MASTER EN SCIENCES MATHÉMATIQUES

"Extension à la programmation convexe conique ""self scaled"" des méthodes de réduction de potentiel utilisées en programmation linéaire."

Vanderbecken, Nathalie

Award date:
1998

Facultés Universitaires Notre-Dame de La Paix
Namur
Faculté des Sciences

Extension à la programmation convexe conique "self-scaled" des méthodes de réduction de potentiel utilisées en programmation linéaire

Mémoire présenté pour l'obtention du grade
de Licencié en Sciences
mathématiques
par

VANDERBECKEN
Nathalie

Promoteur : **J.-J. STRODIOT**

Année académique 1997-1998

Je tiens à remercier Monsieur le Professeur Jean-Jacques Strodiot, promoteur de ce mémoire, pour sa disponibilité et sa contribution à la réalisation de ce travail.

Mes remerciements s'adressent également à tous ceux qui ont eu la gentillesse de consacrer un peu de leur temps et de leur attention à ce mémoire, en particulier Laurent dont les conseils et l'aide m'ont fait gagner un temps considérable.

Je remercie, finalement, mes parents pour m'avoir permis de mener à terme mes études ainsi que mes amis qui m'ont rendu ces quatre années très agréables.

Extension à la programmation convexe conique "self-scaled" des méthodes de réduction de potentiel utilisées en programmation linéaire

VANDERBECKEN Nathalie

Résumé:

L'objectif de ce mémoire est d'élaborer des méthodes de réduction de potentiel à grands pas pour des problèmes de programmation convexe conique dont le cône est "self-scaled". Nous montrons que l'extension à la programmation convexe conique des méthodes de réduction de potentiel primales et primale-duale symétrique utilisées en programmation linéaire, donne lieu à des algorithmes à grands pas lorsque les cônes considérés sont "self-scaled". Notre démarche est la suivante. Nous établissons, dans une première partie, différentes propriétés des cônes "self-scaled". Une seconde partie qui détaille l'extension à la programmation convexe conique "self-scaled" des méthodes de réduction de potentiel utilisées en programmation linéaire, exploite ces propriétés afin de garantir la convergence des extensions obtenues ainsi que la possibilité de progresser à chaque itération d'une grande fraction de la distance séparant l'itéré actuel de la frontière de la région admissible.

Extension to self-scaled conic convexe programming of the potential-reduction methods used in linear programming

VANDERBECKEN Nathalie

Abstract:

The aim of the present work is to elaborate long-step potential-reduction methods for conic convex programming problems when the cone is self-scaled. We show how the restriction to such cones permits to derive primal long-step and symmetric long-step primal-dual algorithms from the potential-reduction methods used in linear programming. In a first part, we establish different properties of the self-scaled cones. The detail of the extension to self-scaled conic convex programming of the potential-reduction methods used in linear programming forms the subject of a second part. This exploits the first established properties in order to guarantee the convergence of the extensions as well as the possibility of taking long step at each iteration.

Mémoire de Licence en Sciences Mathématiques, Unité d'Optimisation.
Juin 1998.
Promoteur: J.-J. STRODIOT.

Table des matières

Introduction

Yu. Nesterov et M.J. Todd étudient dans [1] la résolution d'une extension du problème de programmation linéaire standard au moyen de méthodes de points intérieurs. L'extension envisagée appelée programmation convexe conique, consiste à contraindre la variable primale à appartenir non plus à l'orthant non négatif mais à un cône convexe fermé de l'espace de Hilbert réel considéré.

Des méthodes de points intérieurs applicables à cette extension étaient déjà disponibles dans la littérature. En effet, une généralisation à la programmation convexe d'algorithmes de points intérieurs utilisés en programmation linéaire, a notamment été réalisée par Yu. Nesterov et A. Nemirovskii dans [2]. Ces généralisations permettent la résolution du problème de programmation convexe conique puisque celui-ci est un cas particulier de la programmation convexe. L'intérêt de la démarche menée par Yu. Nesterov et M.J. Todd ne se situe donc pas —du moins, pas totalement— dans l'élaboration de nouvelles méthodes de points intérieurs solutionnant l'extension envisagée. Les méthodes qu'ils décrivent sont effectivement toutes des extensions ou des particularisations d'algorithmes existant déjà respectivement en programmation linéaire et en programmation convexe conique. L'originalité de leur démarche réside, en fait, dans l'introduction d'une hypothèse supplémentaire. Ils supposent "self-scaled" le cône auquel doit appartenir la variable primale. Cette supposition permet, sans réduire excessivement leur applicabilité, une sensible amélioration de la rapidité de convergence des algorithmes résolvant le problème de programmation convexe conique.

Ce mémoire se propose de rendre plus accessible le travail réalisé par Yu. Nesterov et M.J. Todd dans [1].

Il est structuré comme suit.

Une première partie regroupant les quatre premiers chapitres, définit et établit différentes propriétés des cônes "self-scaled".
Le premier chapitre est consacré à la définition du cadre de travail dans lequel le problème de programmation convexe conique a jusqu'à présent été traité.
Dans le second chapitre, nous définissons la classe des barrières et des cônes "self-scaled"

à laquelle nous nous restreindrons. Nous en donnons quelques exemples. Ceux-ci nous permettent alors de juger de la perte de généralité engendrée par l'introduction de l'hypothèse de "self-scaling". Ce deuxième chapitre comprend également l'établissement des propriétés fondamentales des cônes "self-scaled".
Plusieurs mesures de la distance d'un point de l'espace à la frontière d'un cône "self-scaled" sont introduites au troisième chapitre. Sur base des propriétés spécifiques de ces mesures, nous dégageons une borne sur la variation d'une barrière "ν-self-scaled" valable pour une longueur de pas plus grande que celle pour laquelle elle aurait été garantie en l'absence des nouvelles mesures introduites et de l'hypothèse de "self-scaling".
Le quatrième chapitre étudie, quant à lui, le comportement d'une barrière "ν-self-scaled" sur un cône bidimensionnel défini par des directions orthogonales. Cette étude nous permet de déduire une importante conséquence.
Nous pourrons apprécier la pertinence et la portée des propriétés des cônes "self-scaled" ainsi établies dans la seconde partie.

La seconde partie est consacrée à l'élaboration de deux méthodes de réduction de potentiel primales et d'une méthode de réduction de potentiel primale-duale symétrique qui résolvent l'extension envisagée ainsi que son dual dans le cas de la méthode primale-duale. Les méthodes élaborées sont toutes des extensions d'algorithmes utilisés en programmation linéaire. Nous nous efforçons, dans cette partie, de souligner les améliorations rendues possibles par l'introduction de l'hypothèse de "self-scaling".
Le cinquième chapitre achève la définition de notre cadre de travail. Nous y spécifions les problèmes primal et dual traités ainsi que nos hypothèses et leurs conséquences directes. L'étude des projections obliques fait également l'objet de ce chapitre. Cette étude est motivée par le fait que les systèmes linéaires définissant ces projections s'apparentent à ceux que nous aurons à résoudre dans nos algorithmes lors de la détermination des directions de recherche.
Dans le sixième chapitre, nous étendons au problème primal considéré les méthodes de réduction de potentiel primales établies par N.K. Karmarkar (voir [3],[8]) et C.C. Gonzaga (voir [4],[8]) pour la programmation linéaire. Nous explicitons comment la restriction aux cônes "self-scaled" permet d'obtenir des extensions à grands pas.
Nous montrons, dans le dernier chapitre, que la généralisation aux problèmes primal et dual considérés de la méthode de Kojima, Mizuno et Yoshise (voir [5], [8]) établie pour la programmation linéaire, donne lieu à un algorithme primal-dual symétrique à grands pas lorsque le cône primal est supposé "self-scaled".

Cette manière de procéder calque celle utilisée par Yu. Nesterov et M.J. Todd dans [1] si ce n'est que nous restreignons notre étude aux méthodes de réduction de potentiel. Yu. Nesterov et M.J. Todd, quant à eux, soulignent également l'impact de l'introduction de l'hypothèse de "self-scaling" sur une méthode primale de suivi de chemin.

Chapitre 1

Définition du cadre de travail

Ce premier chapitre est consacré à la définition de l'environnement dans lequel le problème de programmation convexe conique a jusqu'à présent été traité. Nous y précisons et, au besoin, y définissons les objets mathématiques nécessaires à la définition de celui-ci. Les hypothèses satisfaites par ces objets y sont également précisées. Nous en profitons pour introduire nos notations et nous y mentionnons aussi des résultats utiles pour la suite.

Nous sommes conscients de l'apparence rébarbative des trois définitions introduites en premier lieu et du fait que des exemples les illustrant auraient été les bienvenus. Cependant, ces définitions ne seront explicitement sollicitées qu'à de très rares occasions dans la suite de ce mémoire. De plus, l'objet de ce dernier est certes la résolution du problème de programmation convexe conique mais dans un cadre de travail restreint. Nous préférons dès lors donner des exemples une fois introduite l'hypothèse supplémentaire que nous considérerons. Le lecteur désirant tout de même disposer d'exemples relatifs aux trois définitions en question, peut se référer aux cas détaillés en 2.2 dont les barrières respectives vérifient ces définitions.

Soit E un espace de Hilbert réel de dimension finie.

Définition 1.1 *Soient a un réel strictement positif, O une partie convexe, ouverte, non vide de E et F une fonction à valeurs réelles définie sur O.*
F est dite a-self-concordante sur O si elle est convexe et de classe C^3 sur O et si elle satisfait, quels que soient x un élément de O et p une direction de E, l'inégalité suivante :

$$|D^3F(x)[p,p,p]| \leq 2a^{-\frac{1}{2}}(D^2F(x)[p,p])^{\frac{3}{2}}$$

où $D^kF(x)[p_1,\ldots,p_k]$ désigne la dérivée directionnelle d'ordre k de F en x dans les directions $p_1,\ldots,p_k$.

Le caractère a-self-concordant de F sur O devient de la forte a-self-concordance lorsque tous les ensembles niveau de F sont fermés.

D'autre part, O est dit être un domaine naturel de F si la suite $(F(x_i))_{i \in I\!N_0}$ converge vers l'infini quelle que soit la suite $(x_i)_{i \in I\!N_0}$ d'éléments de O convergeant vers un point de la frontière de O.

Définition 1.2 *Soient ν un réel positif, C une partie convexe, fermée, à intérieur non vide de E et F une fonction à valeurs réelles définie sur l'intérieur de C.*
Nous définissons, en tout x point intérieur de C, la quantité suivante :

$$\lambda(F,x) \stackrel{def}{=} \begin{cases} \inf\{\lambda \ : \ |DF(x)[p]| \leq \lambda a^{\frac{1}{2}}(D^2F(x)[p,p])^{\frac{1}{2}} \quad \forall p \in E\} \\ \qquad \text{si } \{\lambda \ : \ |DF(x)[p]| \leq \lambda a^{\frac{1}{2}}(D^2F(x)[p,p])^{\frac{1}{2}} \quad \forall p \in E\} \text{ est non vide,} \\ \\ +\infty \qquad \text{sinon} \end{cases}$$

et nous qualifions F de barrière ν-self-concordante pour C si F est une fonction fortement 1-self-concordante sur l'intérieur de C qui vérifie l'inégalité ci-dessous :

$$\sup\{\lambda^2(F,x) \ : \ x \in int\, C\} \leq \nu.$$

Définition 1.3 *Soient ν un réel supérieur ou égal à 1, K un cône convexe, fermé, à intérieur non vide, propre de E et F une fonction à valeurs réelles définie sur l'intérieur de K.*
Nous qualifions F de barrière ν-logarithmiquement homogène pour K si F est une fonction convexe, de classe C^2, dont l'intérieur de K est un domaine naturel et qui satisfait, quels que soient x un point intérieur de K et τ un réel strictement positif, l'égalité suivante :

$$F(\tau x) = F(x) - \nu \ln \tau.$$

Par ailleurs, la fonction F est dite être une barrière ν-normale pour K s'il s'agit d'une barrière 1-self-concordante et ν-logarithmiquement homogène pour K.

Nous désignons l'espace dual de E par E^* et le produit scalaire correspondant à E et à son dual par

$$< s, x >$$

où x et s sont des éléments respectifs de E et de E^*. Nous considérons également K un cône convexe, fermé, à intérieur non vide, propre de E ainsi que F une barrière ν-normale pour K. De plus, nous supposons le cône K pointé et les hessiens de F symétriques et définis positifs. Signalons que la convexité de F garantissait déjà le caractère semi-défini positif de ses hessiens. Par conséquent, l'unique utilité de l'hypothèse suivant laquelle ceux-ci sont définis positifs, est d'assurer leur inversibilité. L'hypothèse de symétrie, couplée au fait que les espaces de Hilbert considérés sont réels, implique, quant à elle, l'égalité entre les hessiens de F et leurs adjoints. Cette propriété de la barrière F ainsi que celles répertoriées dans la proposition suivante seront exploitées à de nombreuses reprises dans la suite de ce mémoire.

Proposition 1.1 *La barrière ν-normale F associée au cône K considéré satisfait, quels que soient x un point intérieur de K et τ un réel strictement positif, les égalités suivantes :*

$$F'(\tau x) = \frac{1}{\tau}F'(x), \qquad F''(\tau x) = \frac{1}{\tau^2}F''(x), \tag{1.1}$$

$$F''(x)x = -F'(x), \qquad F'''(x)[x] = -2F''(x), \tag{1.2}$$

$$< F'(x), x >= -\nu, \tag{1.3}$$

$$< F''(x)x, x >= \nu, \qquad < F'(x), [F''(x)]^{-1}F'(x) >= \nu. \tag{1.4}$$

Explicitons la notation $F'''(x)[x]$ utilisée dans la seconde égalité décrite en (1.2). Etant donnés x un point intérieur de K et p_1 un élément de E, nous désignons par $F'''(x)[p_1]$ l'opérateur linéaire à valeurs dans E^* défini, pour tout p_2 élément de E, par

$$F'''(x)[p_1]p_2 = \frac{d}{dx} < F''(x)p_1, p_2 > .$$

L'intérêt de cette notation est de fournir une expression explicite des dérivées directionnelles d'ordre trois de F. En effet, considérons x un point intérieur de K ainsi que p_1, p_2 et p_3 des éléments de E, et calculons la dérivée directionnelle d'ordre trois de F en x dans les directions p_1, p_2 et p_3. La définition des dérivées directionnelles et le développement de Taylor applicable à F puisque celle-ci est de classe C^3, justifient l'égalité suivante :

$$DF(x)[p_1] =< F'(x), p_1 > .$$

Similairement, nous obtenons, en posant $G(x) =< F'(x), p_1 >$, le développement suivant :

$$\begin{aligned} D^2F(x)[p_1, p_2] &= D(DF(x)[p_1])[p_2] \\ &= DG(x)[p_2] \\ &= < G'(x), p_2 > \\ &= < F''(x)p_1, p_2 > . \end{aligned}$$

Nous posons alors $H(x) =< F''(x)p_1, p_2 >$. La notation introduite et un raisonnement semblable au précédent conduisent à l'expression suivante de la dérivée directionnelle d'ordre trois de F en x dans les directions p_1, p_2 et p_3 :

$$\begin{aligned} D^3F(x)[p_1,p_2,p_3] &= D(D^2F(x)[p_1,p_2])[p_3] \\ &= DH(x)[p_3] \\ &= < H'(x), p_3 > \\ &= < F'''(x)[p_1]p_2, p_3 > . \end{aligned}$$

Preuve de la Proposition 1.1:

Les égalités (1.1) et (1.2) sont obtenues en dérivant respectivement par rapport à x et à τ la relation

$$F(\tau x) = F(x) - \nu \ln \tau$$

satisfaite par F, barrière notamment ν-logarithmiquement homogène pour K.
La Proposition 2.3.4 établie par Yu. Nesterov et A. Nemirovskii dans [2] garantit la satisfaction des égalités (1.3) et (1.4).

□

La définition du cadre de travail dans lequel le problème de programmation convexe conique a jusqu'à présent été traité nécessite encore la définition du cône dual de K et de la conjuguée de F ainsi que l'établissement de certaines de leurs propriétés.

Définition 1.4 *Nous définissons K^*, le cône dual de K, par*

$$K^* \stackrel{def}{=} \{s \in E^* \ : \ < s, x > \geq 0 \qquad \forall x \in K\}.$$

Définition 1.5 *La conjuguée de F, notée F_*, est définie, en tout s point intérieur de K^*, par*

$$F_*(s) \stackrel{def}{=} \max\{- < s, x > -F(x) \ : \ x \in \ int \ K\}.$$

Proposition 1.2 *Le cône dual du cône K considéré est pointé et à intérieur non vide.*

Preuve:

Le caractère pointé de K^* se démontre facilement.
Deux étapes sont nécessaires pour montrer que l'intérieur de K^* est non vide.

1. Nous prouvons que, si l'intérieur de K^* est vide, il existe x un élément non nul de K tel que

$$< s, x >= 0$$

pour tout s point intérieur de K^*. Nous désignons par M le sous-espace vectoriel engendré par K^*. Comme l'intérieur de K^* est supposé vide, nous obtenons l'inégalité stricte suivante :

$$\dim M < \dim E^*.$$

En effet, si ces deux dimensions étaient égales, l'intérieur relatif de K^* coïnciderait avec l'intérieur de K^*. Puisque K^* est convexe et non vide, son intérieur relatif est non vide. Par conséquent, cette coïncidence impliquerait que l'intérieur de K^* est non vide. Le fait que M soit un sous-espace vectoriel propre de E^* garantit l'existence d'un élément non nul de E qui soit orthogonal à tout vecteur de M. Nous montrons que cet élément est le point x de K recherché. L'inclusion de K^* dans M assure l'orthogonalité de cet élément avec tout point intérieur de K^*. D'autre part, cette inclusion, combinée à la fermeture de K, implique les relations suivantes :

$$M^{\perp} \subseteq K^{**} = K$$

et, par suite, l'appartenance de l'élément considéré à K.

2. Nous montrons que l'intérieur de K^* est non vide. Comme $K = K^{**}$, nous pouvons expliciter K de la façon suivante :

$$K = \{x \in E : \quad < s, x > \geq 0 \quad \forall s \in K^*\}.$$

D'autre part, K est supposé pointé. Par conséquent, la propriété suivante est satisfaite :

$$\forall x \neq 0, \in K \qquad -x \notin K.$$

L'expression de K obtenue ci-dessus nous permet de réécrire cette propriété comme suit :

$$\forall x \neq 0, \in K \quad \exists s \in K^* \quad < s, x >> 0.$$

La contraposée de l'implication démontrée en premier lieu permet alors de conclure au résultat escompté.

□

Proposition 1.3 *La conjuguée F_* de la barrière ν-normale F associée au cône K considéré constitue une barrière ν-normale pour K^*, le cône dual de K.*

Preuve:
Ce résultat découle du Théorème 2.4.4 établi par Yu. Nesterov et A. Nemirovskii dans [2].

□

Proposition 1.4 *La conjuguée F_* de la barrière ν-normale F associée au cône K considéré satisfait, quels que soient x et s des points intérieurs respectifs de K et de K^*, les propriétés suivantes :*

$$-F'(x) \in int\ K^*, \qquad -F_*'(s) \in int\ K, \tag{1.5}$$

$$F_*(-F'(x)) =< F'(x), x > -F(x) = -\nu - F(x), \tag{1.6}$$

$$F(-F_*'(s)) = -\nu - F_*(s), \tag{1.7}$$

$$F_*'(-F'(x)) = -x, \qquad F'(-F_*'(s)) = -s, \tag{1.8}$$

$$F''(-F_*'(s)) = [F_*''(s)]^{-1}, \qquad F_*''(-F'(x)) = [F''(x)]^{-1}, \tag{1.9}$$

$$F(x) + F_*(s) \geq -\nu + \nu \ln \nu - \nu \ln < s, x > . \tag{1.10}$$

Les propriétés décrites dans la Proposition 1.4 seront très fréquemment utilisées dans le reste de ce mémoire.

Dans les trois prochains chapitres, nous étudions les répercussions de l'introduction d'une hypothèse supplémentaire sur la barrière ν-normale F associée au cône K considéré.

Chapitre 2

Définition et propriétés fondamentales des barrières et des cônes "self-scaled"

Nous définissons, tout d'abord, la classe des barrières et des cônes à laquelle nous nous restreindrons dans la suite de ce mémoire. Nous jugeons, ensuite, de la perte de généralité engendrée par cette restriction. Finalement, nous établissons les propriétés fondamentales des barrières et cônes appartenant à cette classe.

2.1 Définition des barrières et des cônes "self-scaled"

Définition 2.1 *Etant donnés K un cône pointé à intérieur non vide et F une barrière ν-normale pour K, F est dite être une barrière "ν-self-scaled" pour K si, pour tout x et v points intérieurs de K, nous avons les deux relations suivantes :*

$$\begin{aligned} F''(v)\, x &\in int\ K^* & (2.1) \\ F_*(F''(v)\, x) &= F(x) - 2F(v) - \nu & (2.2) \end{aligned}$$

Si le cône K admet une telle barrière, il est dit être un cône "self-scaled".

Dans le reste de ce mémoire, nous supposerons que la barrière ν-normale F associée au cône K satisfait les relations (2.1) et (2.2). Par conséquent, nous ne travaillerons plus qu'avec une barrière et un cône "self-scaled".

2.2 Caractère non restrictif de la classe des barrières et des cônes "self-scaled"

L'hypothèse de travail suivant laquelle la barrière F considérée est "ν-self-scaled" semble fort restrictive. Dès lors, nous craignons de développer un outil mathématique certes puissant mais d'applicabilité très réduite. Cette éventualité met évidemment en question l'intérêt de ce mémoire. L'enjeu de la présente section est donc considérable puisque celle-ci apporte une réponse à la question de l'appartenance à la classe des cônes "self-scaled", des trois exemples typiques que constituent l'orthant non négatif, le cône des matrices symétriques semi-définies positives et le cône du second ordre.

2.2.1 Orthant non négatif

Proposition 2.1 *La barrière F définie par*

$$F(x) \stackrel{def}{=} -\sum_{i=1}^{n} \ln x_i$$

est une barrière "n-self-scaled" pour l'orthant non négatif noté K et défini par

$$K \stackrel{def}{=} \{x \in I\!R^n \ : \ x_i \geq 0, \quad i = 1 \ldots n\}.$$

<u>**Preuve:**</u>
On vérifie facilement que K est un cône pointé à intérieur non vide. D'autre part, Yu. Nesterov et A. Nemirovskii ont établi dans [2, exemple 3, page 40] le caractère n-normal de la barrière F associée au cône K. Par conséquent, F est une barrière "n-self-scaled" pour K si F satisfait les relations (2.1) et (2.2) pour tout x et v appartenant à $\{x \in I\!R^n \ : \ x_i > 0, \quad i = 1 \ldots n\}$. La vérification de la satisfaction de ces deux relations est triviale lorsque les expressions de K^* et de F_* sont connues. On montre aisément que K^* coïncide avec K. Par ailleurs, les deux assertions ci-dessous :

$$\begin{aligned} F'(x)_i &= \frac{-1}{x_i} \qquad i = 1 \ldots n \\ F''(x) &= \text{diag}(\frac{1}{x_i^2}, \quad i = 1 \ldots n) \end{aligned}$$

conduisent à définir la conjuguée de la barrière F comme suit :

$$F_*(s) = -\sum_{i=1}^{n} \ln s_i - n.$$

□

2.2.2 Cône des matrices symétriques semi-définies positives

Nous désignons par M_n l'espace vectoriel réel des matrices carrées symétriques d'ordre n et nous le munissons du produit scalaire suivant :

$$< S, X >_{M_n} \stackrel{def}{=} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij}\, x_{ij} = \mathrm{Tr}(SX).$$

Proposition 2.2 *La barrière F définie par*

$$F(X) \stackrel{def}{=} -\ln \det X$$

est une barrière "n-self-scaled" pour le cône des matrices symétriques semi-définies positives d'ordre n noté K et défini par

$$K \stackrel{def}{=} \{X \in M_n \ : \ < Xy, y > \geq 0 \qquad \forall y \in I\!R^n\}.$$

Preuve:
On vérifie, à nouveau, facilement que K est un cône pointé à intérieur non vide. D'autre part, la Proposition 5.4.5 établie par Yu. Nesterov et A. Nemirovskii dans [2] garantit le caractère n-normal de la barrière F associée à K. Par conséquent, F est une barrière "n-self-scaled" pour K si F satisfait les relations (2.1) et (2.2) pour tout X et V appartenant à l'intérieur de $\{X \in M_n \ : \ < Xy, y > \geq 0 \quad \forall y \in I\!R^n\}$. La vérification de la satisfaction de ces deux relations nécessite l'établissement de deux propriétés relatives aux dérivées de F ainsi que la connaissance explicite de K^* et de F_*.

1. Nous montrons que K^* coïncide avec K. Nous démontrons, tout d'abord, l'appartenance de X, un élément arbitraire de K, à K^*. Nous devons, donc, par définition de K^*, montrer la positivité de $\mathrm{Tr}(XY)$, quel que soit l'élément Y de K. Dans ce but, nous considérons les décompositions en valeurs propres des matrices X et Y :

$$X = \Lambda_1^T \Gamma_1 \Lambda_1$$
$$Y = \Lambda_2^T \Gamma_2 \Lambda_2$$

où Γ_1 et Γ_2 désignent respectivement les matrices des valeurs propres de X et de Y, et, Λ_1 et Λ_2, les matrices des vecteurs propres de X et de Y. Puisque les matrices des valeurs propres sont diagonales, les propriétés de la trace relatives à de telles matrices nous permettent d'exprimer $\mathrm{Tr}(XY)$ comme suit :

$$\mathrm{Tr}(XY) = \mathrm{Tr}(\Gamma_1 \Lambda_1^T \Lambda_1 \Gamma_2 \Lambda_2^T \Lambda_2).$$

Le caractère orthonormé inhérent aux vecteurs propres justifie l'égalié suivante :

$$\mathrm{Tr}(XY) = \mathrm{Tr}(\Gamma_1\Gamma_2),$$

dont le second membre est effectivement positif puisque les matrices X et Y sont supposées semi-définies positives. Nous avons, de la sorte, démontré l'inclusion de K dans K^*. L'application de ce résultat à K^* et à son dual suffit à démontrer l'inclusion inverse, puisque K est fermé.

2. Nous montrons que $F'(X) = -X^{-1}$. Puisque X appartient à l'intérieur de K, nous pouvons, quel que soit H un élément de M_n, trouver un t strictement positif tel que

$$X + tH \in \text{ int } K.$$

Nous désignons par $\lambda_i, i = 1 \ldots n$, les valeurs propres réelles de la matrice symétrique $X^{-\frac{1}{2}}HX^{-\frac{1}{2}}$. La définition de F, des propriétés élémentaires des déterminants et du logarithme népérien justifient le développement suivant :

$$\begin{aligned}
F(X+tH) &\overset{def}{=} -\ln\det(X+tH)\\
&= -\ln\det[X^{\frac{1}{2}}(I+tX^{-\frac{1}{2}}HX^{-\frac{1}{2}})X^{\frac{1}{2}}]\\
&= \ln(\det[X^{\frac{1}{2}}(I+tX^{-\frac{1}{2}}HX^{-\frac{1}{2}})X^{\frac{1}{2}}])^{-1}\\
&= \ln[\det X^{-\frac{1}{2}}\det(I+tX^{-\frac{1}{2}}HX^{-\frac{1}{2}})^{-1}\det X^{-\frac{1}{2}}]\\
&= \ln\det X^{-1} + \ln\det(I+tX^{-\frac{1}{2}}HX^{-\frac{1}{2}})^{-1}\\
&= \ln\det X^{-1} + \ln[\det(I+tX^{-\frac{1}{2}}HX^{-\frac{1}{2}})]^{-1}\\
&= F(X) + \ln[\prod_{i=1}^{n}(1+t\lambda_i)]^{-1}\\
&= F(X) - \sum_{i=1}^{n}\ln(1+t\lambda_i).
\end{aligned}$$

Remarquons, par ailleurs, que la dernière égalité confirme la convexité de F. Cette confirmation découle de la concavité du logarithme népérien et du résultat établissant l'équivalence entre la convexité d'une fonction définie sur un convexe et la convexité de toute restriction de cette fonction à une droite.
Le développement détaillé ci-dessus conduit à l'égalité suivante :

$$\frac{F(X+tH)-F(X)}{t} = -\sum_{i=1}^{n}\frac{\ln(1+t\lambda_i)}{t}.$$

Dès lors, nous pouvons, en utilisant la règle de l'Hospital, la définition des λ_i et les propriétés de la trace d'un produit de matrices, exprimer la dérivée directionnelle de F en X dans la direction H comme suit :

$$DF(X)[H] \overset{def}{=} \lim_{t\to 0}\frac{F(X+tH)-F(X)}{t}$$

$$\begin{aligned}
&= -\sum_{i=1}^{n} \lim_{t\to 0} \frac{\ln(1+t\lambda_i)}{t} \\
&= -\sum_{i=1}^{n} \lim_{t\to 0} \frac{\lambda_i}{1+t\lambda_i} \\
&= -\sum_{i=1}^{n} \lambda_i \\
&= -\mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) \\
&= \mathrm{Tr}(-X^{-1}H).
\end{aligned}$$

La définition des dérivées directionnelles, couplée au développement de Taylor, conduit à l'égalité suivante :

$$DF(X)[H] =< F'(X), H >_{M_n}$$

dont le second membre équivaut, par définition du produit scalaire considéré, à

$$\mathrm{Tr}(F'(X)H).$$

Nous obtenons, dès lors, la relation suivante :

$$\mathrm{Tr}(-X^{-1}H) = \mathrm{Tr}(F'(X)H).$$

Puisque cette égalité a été établie pour une matrice H arbitraire, nous obtenons bien le résultat escompté.

3. Nous montrons que $F_*(S) = -\ln\det S - n$. Cette expression de la conjuguée de la barrière F est déduite de sa définition. En effet, l'égalité démontrée dans le point précédent permet une détermination aisée des extrema de $- < S, X >_{M_n} - F(X)$ considéré comme une fonction de X. La convexité de $F(X)$ confirmée également dans le point précédent, montre que l'extremum ainsi obtenu est un maximum de la fonction considérée. Dès lors, l'évaluation de celle-ci en ce maximum conduit à l'expression de F_*.

4. Nous montrons que $F''(X)H = X^{-1}HX^{-1}$, quel que soit H appartenant à M_n. Dans ce but, nous établissons, tout d'abord, l'assertion suivante :

$$\frac{d}{dt}[A^{-1}(t)] = -A^{-1}(t)[\frac{d}{dt}A(t)]A^{-1}(t)$$

où A est une application définie sur $I\!R$ et à valeurs dans M_n. Nous considérons deux applications, A et B, définies sur $I\!R$ et à valeurs dans M_n. Nous obtenons, alors, l'égalité suivante :

$$\frac{d}{dt}[A(t)\ B(t)] = [\frac{d}{dt}A(t)]B(t) + A(t)[\frac{d}{dt}B(t)].$$

Dès lors, nous pouvons, si A est inversible, écrire la suite d'égalités ci-dessous :

$$0 = \frac{d}{dt}I = \frac{d}{dt}[A(t)\ A^{-1}(t)] = [\frac{d}{dt}A(t)]A^{-1}(t) + A(t)[\frac{d}{dt}A^{-1}(t)],$$

qui permet d'aboutir au résultat intermédiaire.
Nous démontrons, ensuite, l'objet de ce quatrième point. Nous désignons par G l'application à valeurs dans M_n définie sur l'intérieur de K par

$$G(X) \stackrel{def}{=} -X^{-1}.$$

L'égalité démontrée dans le deuxième point, entraîne, dès lors, que :

$$G(X) = F'(X).$$

Nous considérons H un élément arbitraire de M_n et nous évaluons la dérivée directionnelle de G en X dans la direction H. Les définitions de G et de cette dérivée justifient les égalités ci-dessous :

$$\begin{aligned} DG(X)[H] &= \lim_{t\to 0}\frac{G(X+tH)-G(X)}{t} \\ &= \lim_{t\to 0}\frac{-(X+tH)^{-1}+X^{-1}}{t}. \end{aligned}$$

Nous remarquons que la limite intervenant dans le dernier membre n'est rien d'autre que l'évaluation en 0 de la dérivée de l'application qui, à un réel t, associe la matrice symétrique d'ordre n

$$-(X+tH)^{-1}.$$

Le résultat intermédiaire établi en premier lieu, permet le calcul de cette dérivée. Celle-ci vaut :

$$[(-(X+tH)^{-1})(\frac{d}{dt}(X+tH))(-(X+tH)^{-1})]_{t=0},$$

ce qui, après calcul, se réduit à :

$$X^{-1}HX^{-1}.$$

Puisque $G(X) = F'(X)$, nous obtenons que $DG(X)[H] = F'(X)H$ et donc, nous aboutissons à l'égalité escomptée.

La satisfaction par F des égalités (2.1) et (2.2) se démontre, à présent, aisément grâce à l'égalité demontrée dans le quatrième point et aux propriétés des déterminants.

□

2.2.3 Cône du second ordre

Proposition 2.3 *La barrière F définie par*

$$F(\tau, x) \stackrel{def}{=} -\ln(\tau^2 - \|x\|^2)$$

est une barrière "2-self-scaled" pour le cône du second ordre noté K et défini par

$$K \stackrel{def}{=} \{(\tau, x) \in \mathbb{R}^{n+1} \ : \ \tau \geq \|x\|\},$$

où $\|.\|$ désigne la norme euclidienne sur $\mathbb{R}^n$.

<u>**Preuve:**</u>
Une fois de plus, on vérifie facilement que K est un cône pointé non vide. D'autre part, la Proposition 5.4.3 établie par Yu. Nesterov et A. Nemirovskii dans [2] garantit le caractère 2-normal de la barrière F associée à K. Par conséquent, F est une barrière "2-self-scaled" pour K si F satisfait les relations (2.1) et (2.2) pour tout (η, v) et (τ, x) appartenant à l'intérieur de $\{(\tau, x) \in \mathbb{R}^{n+1} \ : \ \tau \geq \|x\|\}$. La vérification de la satisfaction de ces deux relations nécessite le calcul du gradient et du hessien de F ainsi que la connaissance explicite de K^* et de F_*.

1. Nous montrons que K^* coïncide avec K.
   Nous démontrons, tout d'abord, l'appartenance de (ρ, s), un élément arbitraire de K, à K^*. Nous devons donc, par définition de K^*, montrer la positivité de

   $$\rho\tau + < s, x >,$$

   quel que soit l'élément (τ, x) de K. L'appartenance de (ρ, s) et de (τ, x) à K et l'inégalité de Cauchy-Schwarz permettent respectivement de minorer le premier et le second terme de cette quantité. Nous aboutissons, de la sorte, à l'inégalité suivante:

   $$\rho\tau + < s, x > \geq \|s\|\|x\| - \|s\|\|x\|$$

   qui permet de conclure à l'inclusion de K dans K^*.
   Nous montrons, ensuite, l'inclusion inverse. Nous devons donc, étant donné (ρ, s) un élément arbitraire de K^*, montrer que

   $$\rho \geq \|s\|.$$

   Or, $\|s\| \stackrel{def}{=} \sup_{y \neq 0} \frac{<s,y>}{\|y\|}$. Dès lors, il suffit, pour un y non nul arbitraire, de montrer que

   $$\rho \geq \frac{< s, y >}{\|y\|}.$$

On montre aisément que $(\|y\|, -y)$ est un élément de K. Par conséquent, l'appartenance de (ρ, s) à K^* et la définition de K^* impliquent l'inégalité suivante :

$$\rho\|y\| - < s, y > \geq 0,$$

et donc, le résultat escompté puisque y est distinct de 0.

2. Nous calculons le gradient de F :

$$F'(\tau, x) = \frac{2}{\tau^2 - \|x\|^2} \begin{pmatrix} -\tau \\ x \end{pmatrix}$$

et le hessien de F :

$$F''(\tau, x) = \frac{2}{\tau^2 - \|x\|^2} \begin{pmatrix} -1 & 0 \\ 0 & I_n \end{pmatrix} + \frac{4}{(\tau^2 - \|x\|^2)^2} \begin{pmatrix} \tau^2 & -\tau x^T \\ -\tau x & x x^T \end{pmatrix}$$

3. Nous montrons que $F_*(\rho, s) = -\ln(\rho^2 - \|s\|^2) - 2 + 2\ln 2$.
On vérifie facilement que l'élément (τ, x) de K qui solutionne le système suivant :

$$\begin{aligned} \rho &= \frac{2\tau}{\tau^2 - \|x\|^2} \\ s &= \frac{-2x}{\tau^2 - \|x\|^2}, \end{aligned} \tag{2.3}$$

est le maximum de $-\rho\tau - < s, x > + \ln(\tau^2 - \|x\|^2)$, considéré comme une fonction en (τ, x). Dès lors, l'évaluation de cette fonction en ce maximum conduit à l'expression de F_*.
Nous évaluons, tout d'abord, $\rho\tau + < s, x >$ en (τ, x) solution de (2.3). Nous obtenons ainsi l'égalité suivante :

$$\begin{aligned} \rho\tau + < s, x > &= \frac{2\tau}{\tau^2 - \|x\|^2}\tau - \frac{2}{\tau^2 - \|x\|^2} < x, x > \\ &= 2. \end{aligned}$$

Nous exprimons, ensuite, $\ln(\tau^2 - \|x\|^2)$ évalué en (τ, x) solution de (2.3), en fonction de ρ et de s. Le développement suivant qui exploite la nature de (τ, x) :

$$\begin{aligned} \ln(\rho^2 - \|s\|^2) &= \ln[\frac{4\tau^2}{(\tau^2 - \|x\|^2)^2} - \frac{4|x\|^2}{(\tau^2 - \|x\|^2)^2}] \\ &= \ln(\frac{4}{\tau^2 - \|x\|^2}) \\ &= 2\ln 2 - \ln(\tau^2 - \|x\|^2) \end{aligned}$$

conduit à l'expression ci-dessous :

$$\ln(\tau^2 - \|x\|^2) = -\ln(\rho^2 - \|s\|^2) + 2\ln 2.$$

Les deux évaluations réalisées justifient l'expression de la barrière conjuguée de F.

Nous vérifions, à présent, que F satisfait les égalités (2.1) et (2.2) pour tout (η, v) et (τ, x) appartenant à l'intérieur de $\{(\tau, x) \in \mathbb{R}^{n+1} \ : \ \tau \geq \|x\|\}$.

1. Nous montrons que

$$F''(\eta, v)\begin{pmatrix} \tau \\ x \end{pmatrix}$$

est un point intérieur de K^*. On vérifie aisément, grâce à l'expression du hessien de F établie ci-dessus, l'égalité suivante :

$$F''(\eta, v)\begin{pmatrix} \tau \\ x \end{pmatrix} = \begin{pmatrix} \rho \\ s \end{pmatrix}$$

où

$$\begin{aligned} \rho &= \frac{2[\tau(\eta^2 + \|v\|^2) - 2\eta < v, x >]}{(\eta^2 - \|v\|^2)^2}, \\ s &= \frac{2[(\eta^2 - \|v\|^2)x + 2(< v, x > -\tau\eta)v]}{(\eta^2 - \|v\|^2)^2}. \end{aligned}$$

D'autre part, les expressions de ρ et de s valident le développement suivant :

$$\begin{aligned} &\frac{1}{4}(\eta^2 - \|v\|^2)^4(\rho^2 - \|s\|^2) \\ &= [\tau(\eta^2 + \|v\|^2) - 2\eta < v, x >]^2 - \|(\eta^2 - \|v\|^2)x + 2(< v, x > -\tau\eta)v\|^2 \\ &= \tau^2(\eta^2 + \|v\|^2)^2 - (\eta^2 - \|v\|^2)^2\|x\|^2 - 4\tau^2\eta^2\|v\|^2 \\ &= (\tau^2 - \|x\|^2)(\eta^2 - \|v\|^2)^2. \end{aligned}$$

Nous obtenons, dès lors, l'égalité suivante :

$$\rho^2 - \|s\|^2 = \frac{4(\tau^2 - \|x\|^2)}{(\eta^2 - \|v\|^2)^2}, \tag{2.4}$$

dont le second membre est strictement positif. Cette constatation découle de l'appartenance de (η, v) et de (τ, x) à l'intérieur de K. Par conséquent, l'établissement de la positivité de ρ suffit à garantir que

$$F''(\eta, v)\begin{pmatrix} \tau \\ x \end{pmatrix}$$

est un point intérieur de K^*. Cette positivité se déduit des inégalités et égalité suivantes :

$$\begin{aligned} \rho &\geq \frac{2[\eta^2\tau^{-1}\|x\|^2 + \tau\|v\|^2 - 2\eta < v, x >]}{(\eta^2 - \|v\|^2)^2} \\ &= \frac{2\|\eta\tau^{-\frac{1}{2}}x - \tau^{\frac{1}{2}}v\|^2}{(\eta^2 - \|v\|^2)^2} \\ &\geq 0 \end{aligned}$$

dont la démonstration est aisée.

2. Nous montrons que

$$F_*(F''(\eta, v)\begin{pmatrix} \tau \\ x \end{pmatrix}) = F(\tau, x) - 2F(\eta, v) - 2. \tag{2.5}$$

Les définitions de ρ et de s, et l'expression explicite de F_* permettent de réécrire le premier membre de (2.5) de la façon suivante :

$$-\ln(\rho^2 - \|s\|^2) - 2 + 2\ln 2.$$

On montre alors l'égalité entre cette quantité et le second membre de (2.5) grâce à la relation (2.4), aux propriétés élémentaires du logarithme népérien et à la définition de F.

□

Les Propositions 2.1, 2.2 et 2.3 établissent l'appartenance de l'orthant non négatif, du cône des matrices symétriques semi-définies positives et du cône du second ordre à la classe des cônes "self-scaled". Que cette classe contienne des exemples aussi fondamentaux de la programmation convexe conique, balaie nos craintes d'obtention d'une théorie d'applicabilité réduite et relance, de la sorte, l'intérêt de ce mémoire, un instant mis en doute.

2.3 Etude des propriétés fondamentales des barrières et des cônes "self-scaled"

2.3.1 Application linéaire bijective définie par tout hessien d'une barrière "ν-self-scaled" associée à un cône de ce cône vers son dual

Théorème 2.1

(i.) Quels que soient x et v des points intérieurs de K, nous avons les deux égalités suivantes :

$$\begin{aligned} F_*'(F''(v)\,x) &= [F''(v)]^{-1}\,F'(x) \qquad (2.6)\\ F_*''(F''(v)\,x) &= [F''(v)]^{-1}\,F''(x)\,[F''(v)]^{-1} \qquad (2.7)\end{aligned}$$

(ii.) *Quels que soient v_1 et v_2 deux points intérieurs de K, s'il existe x un point intérieur de K tel que $F''(v_1)\,x = F''(v_2)\,x$,*
alors nous avons que :

$$v_1 = v_2$$

(iii.) *Etant donné v un point intérieur de K fixé, nous avons l'égalité ci-dessous :*

$$K^* = F''(v)\,K$$

Preuve:

(i.) Les égalités (2.6) et (2.7) sont obtenues en dérivant respectivement une fois et deux fois la relation (2.2) satisfaite par F par rapport à x, et en utilisant l'hypothèse de symétrie imposée aux hessiens de F.

(ii.) Nous définissons les deux quantités suivantes :

$$\begin{aligned} s &\stackrel{def}{=} F''(v_1)\,x = F''(v_2)\,x \\ Q &\stackrel{def}{=} \sqrt{F''(x)}. \end{aligned}$$

Nous obtenons, par définition de s et de Q et par application de l'égalité (2.7), le développement suivant pour $i \in \{1,2\}$.

$$\begin{aligned} Q\,F''_*(s)\,Q &= Q\,F''_*(F''(v_i)\,x)\,Q \\ &= Q\,[F''(v_i)]^{-1}\,F''(x)\,[F''(v_i)]^{-1}\,Q \\ &= (\sqrt{F''(x)}\,[F''(v_i)]^{-1}\,\sqrt{F''(x)})^2. \end{aligned}$$

Dès lors, nous avons l'égalité suivante :

$$(Q\,[F''(v_1)]^{-1}\,Q)^2 = (Q\,[F''(v_2)]^{-1}\,Q)^2.$$

Puisque la racine carrée symétrique définie positive d'une matrice carrée définie positive est unique, nous aboutissons à :

$$Q\,[F''(v_1)]^{-1}\,Q = Q\,[F''(v_2)]^{-1}\,Q.$$

Par conséquent, nous obtenons que :

$$F''(v_1) = F''(v_2),$$

qui est une matrice symétrique définie positive que nous désignons par G. La relation (1.2) nous permet d'écrire les deux égalités ci-dessous :

$$F'(v_1) = -G\,v_1$$

$$F'(v_2) = -G\, v_2$$

et donc, nous avons que :

$$< F'(v_1) - F'(v_2), v_1 - v_2 > = \; - < G(v_1 - v_2), v_1 - v_2 > .$$

Puisque F est convexe, nous obtenons l'inégalité suivante :

$$< G(v_1 - v_2), v_1 - v_2 > \; \leq \; 0.$$

Or la matrice G est définie positive. Cette constatation couplée à l'inégalité ci-dessus nous permet de conclure à la nullité de $v_1 - v_2$.

(*iii.*) Nous remarquons, tout d'abord, qu'il est suffisant de prouver l'assertion suivante :

$$\text{int } K^* = F''(v) \text{ int } K.$$

En effet, la topologie nous apprend que cette suffisance est acquise si K et K^* sont fermés et à intérieur non vide. Le cône primal satisfait ces deux propriétés par hypothèse. Quant au cône dual, sa fermeture se démontre aisément et nous avons établi que son intérieur était non vide dans la Proposition 1.2. Dès lors, nous pouvons nous contenter de montrer que int $K^* = F''(v)$ int K. Dans ce but, nous considérons $\hat{s}$ un point intérieur arbitraire de K^* et nous posons $\hat{x} \overset{def}{=} [F''(v)]^{-1}\, \hat{s}$. Puisque F_* est, par la Proposition 1.3, une barrière ν-normale pour K^*, elle est convexe et donc, grâce également à la relation (2.1), nous obtenons, pour tout z point intérieur de K, l'inégalité suivante :

$$F_*(\hat{s}) \geq F_*(F''(v)\, z) + < \hat{s} - F''(v)\, z, F'(F''(v)\, z) > .$$

L'utilisation des relations (2.2), (2.6) et (1.3), de l'hypothèse de symétrie imposée aux hessiens de F et de la définition de $\hat{x}$ nous permet de développer le second membre de cette inégalité comme suit :

$$\begin{aligned}
&F_*(F''(v)\, z) + < \hat{s} - F''(v)\, z, F_*'(F''(v)\, z) > \\
&\quad = \; F(z) - 2F(v) - \nu + < \hat{s} - F''(v)\, z, [F''(v)]^{-1} F'(z) > \\
&\quad = \; F(z) - 2F(v) - \nu + < \hat{s}, [F''(v)]^{-1} F'(z) > - < F''(v)\, z, [F''(v)]^{-1} F'(z) > \\
&\quad = \; F(z) - 2F(v) - \nu + < F'(z), [F''(v)]^{-1} \hat{s} > - < F'(z), [F''(v)]^{-1} F''(v)\, z > \\
&\quad = \; F(z) - 2F(v) - \nu + < F'(z), \hat{x} > - < F'(z), z > \\
&\quad = \; F(z) - 2F(v) + < F'(z), \hat{x} > .
\end{aligned}$$

Nous aboutissons donc à l'inégalité suivante :

$$F_*(\hat{s}) \geq F(z) - 2F(v) + < F'(z), \hat{x} > .$$

Nous posons $s \overset{def}{=} -F'(z)$ et nous obtenons, par la relation (1.8), que :

$$F_*'(s) = F_*'(-F'(z)) = -z.$$

L'inégalité obtenue devient alors la suivante :

$$F_*(\hat{s}) \geq F(-F'_*(s)) - 2F(v) - < s, \hat{x} >$$

dont le second membre peut, aux vues de la relation (1.7), être réécrit comme suit :

$$-\nu - F_*(s) - 2F(v) - < s, \hat{x} > .$$

Par conséquent, nous avons prouvé que, pour tout s point intérieur de K^*, nous avons l'inégalité suivante :

$$F_*(s) + < s, \hat{x} > \geq -F_*(\hat{s}) - 2F(v) - \nu.$$

Or, Yu. Nesterov et A. Nemirovskii ont établi dans [2, Théorèmes 2.4.2 et 2.4.4] que le premier membre de cette inégalité est borné inférieurement si et seulement si $\hat{x}$ appartient à l'intérieur de K. Dès lors, puisque nous sommes parvenus à le borner, nous pouvons conclure à l'appartenance de $\hat{x}$ à l'intérieur de K. Nous avons ainsi montré que, pour $\hat{s}$ un point intérieur de K^*, il existe un unique $\hat{x}$ point intérieur de K tel que $\hat{x} = [F''(v)]^{-1}\,\hat{s}$, c'est-à-dire tel que $\hat{s} = [F''(v)]\,\hat{x}$.

□

Le point (*iii.*) du Théorème 2.1 montre que tout hessien d'une barrière "ν-self-scaled" associée à un cône définit une application linéaire bijective de ce cône vers son dual. Le point (*ii.*) de ce même théorème établit l'unicité des applications ainsi obtenues. En effet, il contraint deux applications distinctes à associer à un même élément de K deux éléments distincts de K^*.

2.3.2 Relations entre les cônes primal et dual relativement aux propriétés des barrières et des cônes "self-scaled"

Proposition 2.4 *Etant donnés K un cône "self-scaled" et F sa barrière "ν-self-scaled" associée, nous avons, pour tout u et s points intérieurs de K^*, les deux relations suivantes :*

$$F''_*(u)\,s \;\in\; int\,K \tag{2.8}$$
$$F(F''_*(u)\,s) \;=\; F_*(s) - 2F_*(u) - \nu \tag{2.9}$$

__Preuve:__

(2.8) Nous posons $v \stackrel{def}{=} -F'_*(u)$ et nous obtenons, par la relation (1.8), que :

$$F'(v) = F'(-F'_*(u)) = -u.$$

La relation (1.5) garantit l'appartenance de v à l'intérieur de K. Par conséquent, l'application du point (*iii.*) du Théorème 2.1 établit l'existence de x un point intérieur de K tel que

$$s \quad = \quad F''(v)\,x. \tag{2.10}$$

Ainsi, nous obtenons, par explicitation de u et emploi de la relation (1.9), la suite d'égalités suivante :

$$F_*''(u)\,s = F_*''(-F'(v))\,s = [F''(v)]^{-1}\,s = x \tag{2.11}$$

qui conclut bien à l'appartenance de $F_*''(u)\,s$ à l'intérieur de K.

(2.9) Nous développons $F_*(s)$ grâce aux relations (2.10), (2.2) et (2.11) :

$$\begin{aligned} F_*(s) &= F_*(F''(v)\,x) \\ &= F(x) - 2F(v) - \nu \\ &= F(F_*''(u)\,s) - 2F(v) - \nu. \end{aligned}$$

Par ailleurs, nous obtenons le développement suivant, en explicitant u et en utilisant la relation (1.6) :

$$\begin{aligned} 2F_*(u) + \nu &= 2F_*(-F'(v)) + \nu \\ &= -2\nu - 2F(v) + \nu \\ &= -2F(v) - \nu. \end{aligned}$$

Par conséquent, nous pouvons réécrire $F_*(s)$ comme suit :

$$F_*(s) = F(F_*''(u)\,s) + 2F_*(u) + \nu,$$

égalité qui est bien celle voulue.

□

Le cône dual K^* est pointé et à intérieur non vide. D'autre part, F_* constitue une barrière ν-normale pour K^*. Ces deux assertions découlent respectivement des Propositions 1.2 et 1.3. Par conséquent, la satisfaction des relations (2.8) et (2.9) par tout u et s points intérieurs de K^*, implique que F_* est une barrière "ν-self-scaled" pour K^*, étant données la Définition 2.1 ainsi que les fermetures de K et de F. La Proposition 2.4 établit dès lors que le fait que F soit une barrière "ν-self-scaled" pour K suffit à entraîner qu'il en soit de même pour F_* relativement à K^*.

Nous observons une symétrie entre les propriétés définissant le caractère "ν-self-scaled" d'un cône primal et celles concluant à ce même caractère pour son dual. Cette symétrie justifie l'établissement des résultats relatifs aux barrières et aux cônes "self-scaled" uniquement pour le cône primal. En effet, les analogues de ces résultats pour le dual auront une forme tout à fait symétrique.

L'hypothèse suivant laquelle la barrière ν-normale F associée au cône K est "ν-self-scaled" valide dès lors les propriétés des barrières et des cônes "self-scaled" non seulement à ce cône et à cette barrière mais également, étant donnée la Proposition 2.4, au cône dual K^* et à sa barrière ν-normale associée F_*.

2.3.3 Existence et unicité d'un point "scaling" pour un couple de points intérieurs de K et K^*

Théorème 2.2 *Quels que soient x et s des points intérieurs respectifs de K et de K^*, il existe un unique ω point intérieur de K, appelé point "scaling", tel que*

$$s = F''(\omega)x.$$

De plus, les points intérieurs x, s et ω considérés satisfont les deux égalités suivantes :

$$\begin{aligned} F'(x) &= F''(\omega)\, F_*'(s) & (2.12)\\ F''(x) &= F''(\omega)\, F_*''(s)\, F''(\omega). & (2.13) \end{aligned}$$

<u>Preuve:</u>

1. Nous montrons l'existence de ω un point intérieur de K tel que

   $$s = F''(\omega)x.$$

   Dans ce but, nous considérons la fonction ϕ définie sur l'intérieur de K par

   $$\phi(v) \stackrel{def}{=} < s, v > - < F'(v), x > .$$

   Nous obtenons alors, par la relation (1.2) et l'hypothèse de symétrie imposée aux hessiens de F, le développement suivant :

   $$\begin{aligned} \phi(v) &= < s, v > + < F''(v)\, v, x > \\ &= < s, v > + < F''(v)\, x, v > . \end{aligned}$$

   Or nous avons l'inégalité suivante :

   $$< F''(v)\, x, v > \geq -F(v) - F_*(F''(v)\, x).$$

En effet, cette inégalité est, par définition de F_*, équivalente à la suivante :

$$< F''(v)\, x, v > \ \geq \ -F(v) - \max_{y \in \mathrm{int}\ K} \{- < F''(v)\, x, y > -F(y)\}$$

qui peut encore être réécrite comme suit :

$$< F''(v)\, x, v > +F(v) \geq \min_{y \in \mathrm{int}\ K} \{< F''(v)\, x, y > +F(y)\}.$$

L'appartenance de v à l'intérieur de K garantit la satisfaction de l'inégalité ci-dessus. Par conséquent, nous aboutissons à la minoration suivante de $\phi(v)$:

$$\phi(v) \geq \ < s, v > -F(v) - F_*(F''(v)\, x)$$

dont le second membre peut, par la relation (2.2), être réexprimé comme suit :

$$< s, v > +F(v) - F(x) - \nu.$$

Nous définissons alors la fonction ψ par

$$\psi(v) \stackrel{def}{=} < s, v > +F(v) - F(x) - \nu$$

et nous montrons que tous les ensembles niveau de cette fonction sont bornés. Pour ce faire, nous utilisons le résultat suivant d'analyse convexe établi par R.T. Rockafellar dans [6, Corollaire 8.7.1] :

Corollaire 1 *Tous les ensembles niveau d'une fonction f convexe, propre et fermée sont bornés si l'un d'entre eux est non vide et borné.*

Puisque ψ est une fonction à valeurs réelles, elle est propre. Le caractère défini positif imposé aux hessiens de F entraîne la stricte convexité de ψ. Par conséquent, cette fonction admet un minimum global unique, solution de $\psi'(v) = 0$. Or $\psi'(v) = s + F'(v)$. Donc, le minimum résout l'équation en v suivante :

$$s = -F'(v)$$

qui peut être solutionnée, grâce à la relation (1.8), comme suit :

$$-F_*'(s) = -F_*'(-F'(v)) = v.$$

Dès lors, l'ensemble niveau suivant :

$$\{v \in \mathrm{int}\ K : \ \psi(v) \leq \psi(-F_*'(s))\}$$

se réduit au singleton $-F_*'(s)$. Nous avons donc trouvé un ensemble niveau non vide et borné. Pour satisfaire complètement les hypothèses du Corollaire 1, il faut encore vérifier que ψ est fermée. Puisque cette fonction est convexe et propre, il suffit, pour prouver sa fermeture, de montrer qu'elle est semi-continue inférieurement. Or cette fonction est continue étant données les continuités du produit scalaire et de F, barrière entre autres ν-logarithmiquement homogène. L'application du Corollaire 1 conclut alors au caractère borné de tous les ensembles niveau de ψ. Ce résultat entraîne que tous les ensembles niveau de ϕ sont bornés puisque ψ minore ϕ. D'autre part, la fonction ϕ est continue. Cette propriété découle de la continuité du produit scalaire et du fait que F est une fonction de classe C^2 puisqu'elle est notamment une barrière ν-logarithmiquement homogène. Cette caractéristique de ϕ implique que tous ses ensembles niveau sont fermés. Par conséquent, ϕ atteint sa borne inférieure puisqu'il s'agit d'une fonction semi-continue inférieurement sur des compacts. Le minimum en question est obtenu en annulant la dérivée première de ϕ et solutionne donc l'équation en ω suivante :

$$s = F''(\omega)\, x.$$

Cette équation admet, par le point ($ii.$) du Théorème 2.1, une solution unique qui est le point "scaling" recherché.

2. Les égalités (2.12) et (2.13) s'obtiennent respectivement sur base des égalités (2.6) et (2.7) appliquées à x et à ω, et dans lesquelles on remplace $F''(\omega)\, x$ par s, substitution légitimée par le résultat démontré ci-dessus.

$\square$

Le Théorème 2.1 a établi que tout point intérieur d'un cône "self-scaled" permet, via l'évaluation en lui du hessien de la barrière "ν-self-scaled" associée, de définir une application linéaire bijective unique de ce cône vers son dual. Le Théorème 2.2 garantit, quant à lui, que, parmi toutes les applications ainsi définies, il en existe une et une seule qui associe un point intérieur de K fixé à un point intérieur de K^* fixé.

Le Théorème 2.2 établit donc, pour tout couple de points intérieurs de K et de K^*, l'existence et l'unicité d'un point "scaling" mais ne fournit pas une méthode de détermination de celui-ci. Toutefois, le calcul du point "scaling" associé à un couple donné de points intérieurs de K et de K^*, est, dans la plupart des cas, assez direct. Ainsi, on vérifie facilement que $\omega = [\mathrm{diag}(x)]^{\frac{1}{2}}[\mathrm{diag}(s)]^{-\frac{1}{2}}e$, où e désigne le vecteur de $I\!R^n$ dont toutes les composantes valent 1, constitue le point "scaling" associé au couple (x, s) de points intérieurs de l'orthant non négatif. D'autre part, la propriété établie dans le quatrième point de la preuve de la Proposition 2.2 permet de montrer que $\Omega = X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}}$ constitue le point "scaling" associé au couple (X, S) de points intérieurs du cône des matrices symétriques semi-définies positives.

Les Théorèmes 2.1 et 2.2 justifient la désignation de barrières et de cônes "self-scaled". La barrière admise par un cône "self-scaled" permet, effectivement, d'associer tout point de ce cône à un élément quelconque de son dual. Cette propriété explique le suffixe *"scaled"*. La notion d'autosuffisance inspirée par le préfixe *"self"* provient du fait qu'il suffit de considérer les points intérieurs du cône "self-scaled" pour envisager toutes les associations possibles entre un point de ce cône et un élément de son dual. Malgré qu'une traduction française de ce terme et de la désignation de point "scaling" aurait été plus agréable à la lecture, nous avons préféré conserver la terminologie anglaise faute d'appellations françaises équivalentes.

Par ailleurs, comme nous l'avons mentionné, la Proposition 2.4 valide l'application du Théorème 2.2 au cône dual K^* et à sa barrière ν-normale associée F_*. Nous obtenons ainsi le résultat suivant :

Corollaire 2.1 *Quels que soient s et x des points intérieurs respectifs de K^* et de K, il existe un unique t point intérieur de K^*, appelé point "scaling", tel que*

$$x = F_*''(t)\, s.$$

De plus, les points intérieurs s, x et t considérés satisfont les deux égalités suivantes :

$$\begin{aligned} F_*'(x) &= F_*''(t)\, F'(x) && (2.14)\\ F_*''(s) &= F_*''(t)\, F''(x)\, F_*''(t). && (2.15) \end{aligned}$$

La démonstration du Corollaire 2.1 découle de l'application du Théorème 2.2 à x et à s et de la relation (1.9) qui conduit à définir t comme suit :

$$t \stackrel{def}{=} -F'(\omega).$$

D'autre part, la relation (1.8) nous permet d'exprimer ω en fonction du t ainsi défini de la façon suivante :

$$\omega = -F_*'(-F'(\omega)) = -F_*'(t).$$

Nous disposons de la sorte d'une illustration de la symétrie régnant entre le cône primal K et le cône dual K^* relativement aux propriétés des barrières et des cônes "self-scaled", symétrie que nous avions annoncée.

2.3.4 Propriétés des dérivées des barrières "ν-self-scaled"

Corollaire 2.2

($i.$) *Quels que soient v et x des points intérieurs de K, l'opérateur $F'''(x)[v]$ est semi-défini négatif.*

($ii.$) *Quels que soient x un point intérieur de K et p une direction de E telle que $x+p$ appartient à K, nous avons l'inégalité suivante :*

$$F'''(x)[p] \leq 2F''(x).$$

Précisons que l'inégalité décrite dans le point ($ii.$) est relative au cône des opérateurs semi-définis positifs. Une telle inégalité signifie que l'opérateur obtenu en soustrayant le premier membre du second est semi-défini positif.

La preuve du Corollaire 2.2 se décompose en deux étapes. D'une part, nous établissons un résultat qui permettra de démontrer le point ($i.$) du Corollaire 2.2 ainsi que le Lemme 4.1. D'autre part, nous prouvons le Corollaire 2.2.

Lemme 2.1 *Etant donné v un élément de K fixé, nous considérons la fonction g définie sur l'intérieur de K par*

$$g(x) \stackrel{def}{=} < -F'(x), v > .$$

Nous avons alors que cette fonction est convexe.

<u>Preuve du Lemme 2.1:</u>

Nous envisageons, tout d'abord, le cas où v appartient à l'intérieur de K. Nous devons montrer que, pour tout x et z points intérieurs de K, nous avons l'inégalité suivante :

$$g(x) \geq g(z) + < g'(z), x - z >$$

qui peut, par explicitation de g et de sa dérivée première, se réécrire comme suit :

$$< -F'(x), v > \; \geq \; - < F'(z), v > - < F''(z)\, v, x - z > .$$

Nous remarquons que démontrer cette inégalité revient à montrer que la fonction en x définie par

$$- < F'(x), v > + < F''(z)\, v, x > \tag{2.16}$$

atteint son minimum en x égal à z. Or, dans la preuve du Théorème 2.2, nous avons démontré que la fonction en x définie par

$$- < F'(x), v > + < s, x >$$

atteint son minimum quel que soit s un point intérieur de K^* et que ce minimum est l'unique solution de l'équation en x suivante:

$$s = F''(x)\, v.$$

Dès lors, il suffit, pour montrer que la fonction décrite en (2.16) atteint son minimum en z, de particulariser ce résultat à $s = F''(z)\, v$. Remarquons que la relation (2.1) satisfaite par F valide cette particularisation. Nous obtenons de la sorte que l'unique minimum de la fonction définie en (2.16) solutionne l'équation en x suivante:

$$F''(z)\, v = F''(x)\, v.$$

Le point $(ii.)$ du Théorème 2.1 implique que x égale z. Cette égalité permet de conclure à la convexité de g.
Le caractère fermé de K, la convexité de g déjà acquise pour v un point intérieur de K, la continuité du produit scalaire et le passage à la limite dans les inégalités permettent la démonstration de la convexité de g dans le cas où v est un élément de la frontière de K.

□

<u>Preuve du Corollaire 2.2:</u>

$(i.)$ Le Lemme 2.1 a établi la convexité de la fonction

$$g(x) \stackrel{def}{=} - < F'(x), v >$$

quels que soient x un point intérieur de K et v un élément de K. Puisque F est une barrière ν-normale pour K, elle est 1-self-concordante sur l'intérieur de K et donc, de classe C^3 sur ce même ensemble. Par conséquent, la fonction g est de classe C^2 et sa convexité est dès lors équivalente au caractère semi-défini positif de son hessien en tout point de son domaine de définition. Comme $g(x)$ n'est en fait rien d'autre que l'opposé de la dérivée directionnelle de F en x dans la direction v, nous avons que le hessien de g en x s'apparente à

$$-F'''(x)[v].$$

Par conséquent, la convexité de g conclut au caractère semi-défini négatif de l'opérateur $F'''(x)[v]$, quels que soient v un élément de K et x un point intérieur de K.

$(ii.)$ Nous désignons par z l'élément de K défini par

$$z \stackrel{def}{=} x + p.$$

Nous avons alors, quel que soit v un élément de E, que:

$$\begin{aligned} < F'''(x)[p]\, v, v > &= < F'''(x)[x + p - x]\, v, v > \\ &= < F'''(x)\,[z] v, v > - < F'''(x)[x]\, v, v > . \end{aligned}$$

Le point $(i.)$ établit la négativité de $< F'''(x)[z]\, v, v >$. D'autre part, la relation (1.2) permet d'exprimer $F'''(x)[x]$ comme $-2F''(x)$. Par conséquent, nous obtenons, quel que soit v appartenant à E, l'inégalité suivante:

$$< F'''(x)[p]\, v, v > \leq < 2F''(x)\, v, v >$$

qui est équivalente à

$$< (2F''(x) - F'''(x)[p])\, v, v > \geq 0.$$

Cette dernière relation établit le caractère semi-défini positif de l'opérateur $2F''(x) - F'''(x)[p]$.

□

Dans ce deuxième chapitre, nous avons notamment décrit les propriétés fondamentales des barrières et des cônes "self-scaled". Les Théorèmes 2.1 et 2.2 établissent les deux principaux résultats. La Proposition 2.4, couplée à l'hypothèse suivant laquelle F constitue une barrière "ν-self-scaled" pour le cône K, étend la validité des résultats obtenus pour le cône primal au cône dual. Le Corollaire 2.2 fournit, quant à lui, deux propriétés importantes des dérivées des barrières "ν-self-scaled".

Les deux chapitres suivants étudient et exploitent des propriétés plus spécifiques des barrières et des cônes "self-scaled" en vue de dégager des théorèmes dont nous pourrons apprécier la pertinence et la portée lors de la détermination des méthodes de réduction de potentiel.

Chapitre 3

Distance à la frontière d'un cône "self-scaled"

Nous débutons ce chapitre par la définition et l'établissement des principales propriétés de plusieurs mesures de la distance d'un point de l'espace à la frontière d'un cône "self-scaled" relativement à un point intérieur de celui-ci ou de son dual.

Ces nouveaux outils nous permettent de dégager une borne sur la variation de la barrière "ν-self-scaled" F associée au cône K considéré, valable pour une longueur de pas plus grande que celle pour laquelle elle aurait été garantie en l'absence de l'hypothèse de "self-scaling". Nous soulignons comment l'introduction des nouvelles distances à la frontière de K et la restriction aux barrières "ν-self-scaled" rendent possible l'obtention d'une telle borne.

3.1 Définition de distances à la frontière d'un cône "self-scaled"

3.1.1 Mesures primales

Définition 3.1 *Soient x un point intérieur de K fixé et p un élément de E. Nous définissons une première mesure de la distance de p à la frontière de K relativement à x par*

$$\sigma_x(p) \stackrel{def}{=} \frac{1}{\sup\{\alpha : x - \alpha p \in K\}}.$$

Le peu de maniabilité de cette définition nous incite à considérer l'égalité suivante :

$$\sigma_x(p) \quad = \quad \min\{\beta \geq 0 : \beta x - p \in K\} \tag{3.1}$$

dont la démonstration est immédiate.

Par ailleurs, de simples manipulations de celle-ci permettent d'obtenir les propriétés suivantes :

- $0 \leq \sigma_x(p) < +\infty$ (3.2)
- $x - \alpha p \in \text{ int } K \qquad \forall \alpha \in [0, \frac{1}{\sigma_x(p)})$ (3.3)

  où, par convention, $\frac{1}{\sigma_x(p)} = +\infty \quad \text{si} \quad \sigma_x(p) = 0$
- $\forall v \in K : \quad \sigma_x(x - v) \leq 1 \quad \text{et} \quad \sigma_x(-v) = 0$ (3.4)
- $\sigma_x(-v) = 0 \quad \Leftrightarrow \quad v \in K$ (3.5)
- $\sigma_x(x) = 1 \quad \text{et} \quad \sigma_x(-x) = 0$ (3.6)
- $\forall \lambda > 0 : \quad \sigma_x(\lambda p) = \lambda \sigma_x(p) = \sigma_{\frac{x}{\lambda}}(p)$ (3.7)

Définition 3.2 *Soient x un point intérieur de K fixé, p un élément de E et q un élément de E^*. Nous définissons la norme primale relative à x par*

$$\|p\|_x \stackrel{def}{=} < F''(x)p, p >^{\frac{1}{2}}$$

et la norme duale relative à x par

$$\|q\|_x^* \stackrel{def}{=} < q, [F''(x)]^{-1}q >^{\frac{1}{2}} .$$

Définition 3.3 *Soient x un point intérieur de K fixé et p un élément de E. Nous définissons une seconde mesure de la distance de p à la frontière de K relativement à x par*

$$|p|_x \stackrel{def}{=} \max\{\sigma_x(p), \sigma_x(-p)\}.$$

Ces deux dernières définitions mènent aux inégalités suivantes :

Proposition 3.1 *Quels que soient x un point intérieur de K, p et v des éléments de E et q un élément de E^*, nous avons les propriétés suivantes :*

- $< F''(x)v, p > \leq \|v\|_x \|p\|_x \quad \textit{et} \quad < q, p > \leq \|q\|_x^* \|p\|_x$ (3.8)
- $< -F'(x), p > \leq \|x\|_x \|p\|_x = \sqrt{\nu} \|p\|_x$ (3.9)
- $< -F'(x), p > \leq \nu \sigma_x(p)$ (3.10)
- $\sigma_x(p) \leq \|p\|_x \quad \sigma_x(-p) \leq \|p\|_x \quad |p|_x \leq \|p\|_x$ (3.11)

Preuve:

(3.8) La seconde inégalité est une simple généralisation de l'inégalité de Cauchy-Schwarz aux espaces duaux. La première, quant à elle, s'obtient aisément par application de la seconde et le fait que $\|F''(x)v\|_x^*$ égale $\|v\|_x$.

(3.9) Dans un premier temps, nous utilisons la relation (1.2). Nous appliquons au produit scalaire ainsi obtenu la première inégalité de (3.8) avec l'identification $v = x$. Nous concluons alors par l'explicitation de $\|x\|_x$ grâce à (1.4).

(3.10) Les définitions de K^* et de $\sigma_x(p)$, couplées à la relation (1.5), impliquent l'inégalité suivante :

$$< -F'(x), x - \alpha p > \geq 0 \qquad \forall \alpha \in [0, \frac{1}{\sigma_x(p)}].$$

La première des propriétés décrites en (1.2) et la linéarité du produit scalaire nous permettent de réécrire celle-ci comme suit :

$$< F''(x)x, x > -\alpha < F''(x)x, p > \geq 0 \qquad \forall \alpha \in [0, \frac{1}{\sigma_x(p)}].$$

En appliquant la relation (1.4) au premier produit scalaire et en utilisant la relation (1.2) pour expliciter $F''(x)x$ dans le second, nous aboutissons à la relation ci-dessous :

$$\nu - \alpha < -F'(x), p > \geq 0 \qquad \forall \alpha \in [0, \frac{1}{\sigma_x(p)}]$$

et, par suite, à l'inégalité suivante :

$$\alpha < -F'(x), p > \leq \nu \qquad \forall \alpha \in [0, \frac{1}{\sigma_x(p)}].$$

La particularisation de cette dernière à $\alpha = \frac{1}{\sigma_x(p)}$ nous conduit à la relation suivante :

$$\frac{1}{\sigma_x(p)} < -F'(x), p > \leq \nu$$

qui conclut bien au résultat escompté, étant donné la stricte positivité de $\sigma_x(p)$.

(3.11) Nous rappelons un résultat obtenu par Yu. Nesterov et A. Nemirovskii dans [2,Théorème 2.1.1] :

Théorème 1 *Soient C une partie convexe, ouverte, non vide de E, x un élément de C et F une fonction fortement ν-self-concordante sur C.*
Alors chaque ellipsoïde de Dinkin de F centré en x et de rayon strictement inférieur à 1 est contenu dans C, ce qui mathématiquement est équivalent à

$$r < 1 \quad \Rightarrow \quad \{v : \|x - v\|_{x,F} \leq r\} \subset C$$

où $\|p\|_{x,F} \stackrel{def}{=} (\frac{1}{\nu} D^2 F(x)[p,p])^{\frac{1}{2}}$, *pour tout p appartenant à E.*

Puisque la fonction F que nous considérons est une barrière entre autres 1-self-concordante pour K, elle est fortement 1-self-concordante sur l'intérieur de K. D'autre part, cet intérieur est une partie ouverte, convexe, non vide de E à laquelle x est supposé appartenir. Dès lors, les hypothèses du Théorème 1 sont satisfaites. Par ailleurs, nos identifications entraînent que la norme $\|.\|_{x,F}$ définie dans l'énoncé de Théorème 1 s'apparente à la norme $\|.\|_x$ décrite dans la Définition 3.2. Par conséquent, l'application du Théorème 1 fournit l'implication suivante :

$$r < 1 \quad \Rightarrow \quad \{v : \|v - x\|_x \leq r\} \subset \text{ int } K \subseteq K$$

qui peut être étendue comme suit :

$$r \leq 1 \quad \Rightarrow \quad \{v : \|v - x\|_x \leq r\} \subseteq K$$

par simple utilisation du caractère fermé de K.

Nous exploitons l'inclusion ainsi obtenue afin de démontrer l'inégalité $\sigma_x(p) \leq \|p\|_x$. Dans un premier temps, nous envisageons le cas où $\|p\|_x$ est nulle, autrement dit le cas où p est le vecteur nul. Cette hypothèse sur p implique l'annulation de $\sigma_x(p)$. L'inégalité à démontrer est dès lors triviale. Deuxièmement, nous considérons l'alternative où $\|p\|_x$ vaut un réel r strictement positif. Nous avons alors l'égalité suivante :

$$\|\frac{p}{r}\|_x = 1$$

et donc, la relation ci-dessous :

$$\|x - (x - \frac{p}{r})\|_x = 1.$$

L'inclusion acquise par application du Théorème 1 permet de conclure à l'appartenance de $x - \frac{p}{r}$ à K et, par suite, à celle de $rx - p$ à ce même cône. Cette dernière appartenance implique bien la satisfaction de la relation suivante :

$$\sigma_x(p) \leq r = \|p\|_x.$$

La preuve de la seconde inégalité à démontrer utilise un argument similaire. Cette fois, v est identifié à $x + \frac{p}{r}$.
Quant à la troisième inégalité, elle découle directement des deux précédentes compte tenu de la définition de $|p|_x$.

$\square$

3.1.2 Mesures duales

Définition 3.4 *Soient x et s des points intérieurs respectifs de K et de K^* fixés et q un élément de E^*. Nous définissons deux mesures de la distance de q à la frontière de K^*, l'une relative à s*

$$\sigma_s^*(q) \stackrel{def}{=} \frac{1}{\sup\{\alpha : s - \alpha q \in K^*\}}$$

et l'autre relative à x

$$\sigma_x(q)^* \stackrel{def}{=} \sigma^*_{-F'(x)}(q).$$

Remarquons que $\sigma_s^*(q)$ est l'analogue de $\sigma_x(p)$ pour le cône dual. A nouveau, nous préférons à cette définition peu maniable l'égalité suivante :

$$\sigma_s^*(q) \quad = \quad \min\{\beta \geq 0 : \beta s - q \in K^*\}$$

dont la démonstration est immédiate. Notons également que la définition de $\sigma_x(q)^*$ a bien un sens puisque la première des appartenances décrites en (1.5) garantit que $-F'(x)$ est un point intérieur de K^*.

Définition 3.5 *Soient s un point intérieur de K^* fixé, p un élément de E et q un élément de E^*. Nous définissons la norme primale relative à s par*

$$\|p\|_s \stackrel{def}{=} < [F_*''(s)]^{-1}\ p, p >^{\frac{1}{2}}$$

et la norme duale relative à s par

$$\|q\|_s^* \stackrel{def}{=} < q, F_*''(s)\ q >^{\frac{1}{2}} .$$

Définition 3.6 *Soient x et s des points intérieurs respectifs de K et de K^* fixés et q un élément de E^*. Nous définissons deux nouvelles mesures de la distance de q à la frontière de K^*, l'une relative à s*

$$|q|_s^* \stackrel{def}{=} \max\{\sigma_s^*(q), \sigma_s^*(-q)\}$$

et l'autre relative à x

$$|q|_x^* \stackrel{def}{=} \max\{\sigma_x(q)^*, \sigma_x(-q)^*\}.$$

Soulignons une nouvelle fois l'analogie règnant entre les cônes primal et dual.

3.2 Bornes sur les hessiens et sur la variation d'une barrière "ν-self-scaled"

3.2.1 Supériorité des barrières "ν-self-scaled" sur les barrières self-concordantes

Théorème 3.1 *Quels que soient x un point intérieur de K fixé et p un élément de E, nous avons, pour tout α appartenant à $[0, \frac{1}{\sigma_x(p)})$, la double inégalité suivante :*

$$\frac{1}{(1+\alpha\sigma_x(-p))^2}\, F''(x) \;\leq\; F''(x-\alpha p) \;\leq\; \frac{1}{(1-\alpha\sigma_x(p))^2}\, F''(x). \qquad (3.12)$$

La preuve du Théorème 3.1 se décompose en deux étapes. D'une part, nous établissons un résultat qui permettra de démontrer le Théorème 3.1 ainsi que le Théorème 3.2. D'autre part, nous prouvons la double inégalité (3.12).

Lemme 3.1 *Quels que soient x un point intérieur de K fixé et p un élément de E, nous définissons $\bar{p}$ de la façon suivante :*

$$\bar{p} \stackrel{def}{=} \begin{cases} \frac{-p}{\sigma_x(p)} & \text{si } \sigma_x(p) > 0 \\ \frac{-p}{\sigma} \quad \text{où } \sigma \text{ est un réel strictement positif tel que } x - \frac{p}{\sigma} \text{ appartient à } K & \text{sinon.} \end{cases}$$

Etant donné v une direction de E fixée, nous considérons la fonction définie par

$$\phi(\beta) \stackrel{def}{=} < F''(x(\beta))\, v, v >,$$

où $x(\beta) \stackrel{def}{=} x + \beta\bar{p}$, $\beta \in [0,1)$.
Alors, nous avons que cette fonction satisfait l'inégalité suivante :

$$\phi(\beta) \;\leq\; \frac{\phi(0)}{(1-\beta)^2}.$$

<u>**Preuve du Lemme 3.1:**</u>

1. Nous montrons que $\phi'(\beta) \leq \frac{2}{1-\beta}\, \phi(\beta)$. Nous remarquons que

$$\phi'(\beta) = < F'''(x(\beta))\, [\bar{p}]\, v, v > .$$

Par conséquent, nous voulons montrer l'inégalité suivante :

$$< F'''(x(\beta))\ [\bar{p}]\ v, v > \leq \frac{2}{1-\beta} \ < F''(x(\beta))\ v, v > .$$

Pour ce faire, nous utilisons le point (*ii.*) du Corollaire 2.2 avec les identifications $x = x(\beta)$ et $p = (1-\beta)\ \bar{p}$. On démontre aisément que ces dernières vérifient les hypothèses du point (*ii.*) du Corollaire 2.2. Il suffit d'expliciter $x(\beta)$ et $\bar{p}$, et d'utiliser les définitions de $\sigma_x(p)$ et de σ. L'application du point (*ii.*) du Corollaire 2.2 nous fournit alors l'inégalité suivante :

$$F'''(x(\beta))\ [(1-\beta)\ \bar{p}] \leq 2\, F''(x(\beta))$$

qui est équivalente à

$$(1-\beta)\ F'''(x(\beta))\ [\bar{p}] \leq 2\, F''(x(\beta)).$$

Par définition des inégalités relatives au cône des opérateurs semi-définis positifs, nous avons que l'opérateur $2\, F''(x(\beta)) - (1-\beta)\, F'''(x(\beta))\ [\bar{p}]$ est semi défini-positif. L'explicitation de cette caractéristique, combinée à la linéarité du produit scalaire et à l'appartenance de β à $[0,1)$, nous permet d'aboutir, pour tout v appartenant à E, à la relation suivante :

$$< F'''(x(\beta))\ [\bar{p}]\ v, v > \leq \frac{1}{1-\beta} \ < 2\, F''(x(\beta))\ v, v > .$$

La particularisation à la direction v fixée de cette inégalité vraie quelle que soit l'élément v de E considéré conduit au résultat intermédiaire attendu.

2. Pour obtenir le résultat final, nous posons $\psi(\beta) \stackrel{def}{=} \ln \phi(\beta)$. L'inégalité que nous venons de démontrer devient alors la suivante :

$$\psi'(\beta) \leq \frac{2}{1-\beta}.$$

Par intégration de 0 à β des deux membres de celle-ci et par monotonie de l'intégrale, nous obtenons la relation ci-dessous :

$$\psi(\beta) \leq \psi(0) - 2\ln(1-\beta).$$

Nous aboutissons au résultat escompté, à savoir :

$$\phi(\beta) \leq \frac{\phi(0)}{(1-\beta)^2},$$

en utilisant la définition de $\psi(\beta)$ et des propriétés élémentaires du logarithme népérien.

$\square$

Preuve du Théorème 3.1:

1. Nous démontrons la seconde inégalité de (3.12) dans le cas où $\sigma_x(p)$ est strictement positif. Nous choisissons arbitrairement une direction v de E et nous définissons $\bar{p}$ comme décrit dans l'énoncé du Lemme 3.1. Par application de celui-ci, nous obtenons l'inégalité suivante :

$$\phi(\beta) \;\leq\; \frac{\phi(0)}{(1-\beta)^2},$$

où $\phi(\beta) \stackrel{def}{=} < F''(x(\beta))\, v, v >$ et $x(\beta) \stackrel{def}{=} x + \beta \bar{p}$, $\beta \in [0,1)$. Nous particularisons ensuite cette inégalité à $\beta = \alpha\, \sigma_x(p)$. Remarquons que l'appartenance de $\alpha\, \sigma_x(p)$ à $[0{,}1)$ est garantie par l'hypothèse suivant laquelle α appartient à $[0, \frac{1}{\sigma_x(p)})$. Cette particularisation et l'explicitation de $\phi(\beta)$, de $x(\beta)$ et de $\bar{p}$ conduisent au développement suivant :

$$\begin{aligned} \phi(\alpha\, \sigma_x(p)) \;&\leq\; \frac{\phi(0)}{(1-\alpha\, \sigma_x(p))^2} \\ < F''(x(\alpha\, \sigma_x(p)))\, v, v > \;&\leq\; \frac{< F''(x(0))\, v, v >}{(1-\alpha\, \sigma_x(p))^2} \\ < F''(x + \alpha\, \sigma_x(p)\, \bar{p})\, v, v > \;&\leq\; \frac{< F''(x)\, v, v >}{(1-\alpha\, \sigma_x(p))^2} \\ < F''(x - \alpha\, p)\, v, v > \;&\leq\; \frac{< F''(x)\, v, v >}{(1-\alpha\, \sigma_x(p))^2}. \end{aligned}$$

Puisque cette inégalité à été démontrée pour une direction v de E arbitrairement choisie, elle est vraie quel que soit l'élément v de E considéré. Dès lors, nous obtenons, par définition des inégalités relatives au cône des opérateurs semi-définis positifs, la relation ci-dessous :

$$F''(x - \alpha\, p) \leq \frac{F''(x)}{(1-\alpha\, \sigma_x(p))^2}$$

qui est bien le résultat attendu.

2. Nous démontrons la seconde inégalité de (3.12) dans le cas où $\sigma_x(p)$ est nul. Comme ci-dessus, nous choisissons arbitrairement une direction v de E et nous définissons $\bar{p}$ comme décrit dans l'énoncé du Lemme 3.1. Par application de celui-ci, nous obtenons, à nouveau, l'inégalité suivante :

$$\phi(\beta) \;\leq\; \frac{\phi(0)}{(1-\beta)^2},$$

où $x(\beta) \stackrel{def}{=} x + \beta\bar{p}$ et $\phi(\beta) \stackrel{def}{=} < F''(x(\beta))\, v, v >$, $\beta \in [0,1)$. Nous particularisons ensuite cette inégalité à $\beta = \alpha\,\sigma$. Remarquons que le respect de l'appartenance de β à $[0,1)$ nécessite, cette fois, l'introduction d'une contrainte supplémentaire sur σ : celui-ci doit effectivement être choisi tel que $\sigma < \frac{1}{\alpha}$. Cette particularisation et l'explicitation de $\phi(\beta)$, de $x(\beta)$ et de $\bar{p}$ conduisent au développement suivant :

$$\begin{aligned}
\phi(\alpha\,\sigma) &\leq \frac{\phi(0)}{(1-\alpha\,\sigma)^2} \\
< F''(x(\alpha\,\sigma))\, v, v > &\leq \frac{< F''(x(0))\, v, v >}{(1-\alpha\,\sigma)^2} \\
< F''(x + \alpha\,\sigma\,\bar{p})\, v, v > &\leq \frac{< F''(x)\, v, v >}{(1-\alpha\,\sigma)^2} \\
< F''(x - \alpha\,p)\, v, v > &\leq \frac{< F''(x)\, v, v >}{(1-\alpha\,\sigma)^2}.
\end{aligned}$$

A nouveau, puisque cette inégalité à été démontrée pour une direction v de E arbitrairement choisie, elle est vraie quel que soit l'élément v de E considéré. Dès lors, nous obtenons, par définition des inégalités relatives au cône des opérateurs semi-définis positifs, l'inégalité ci-dessous :

$$F''(x - \alpha\,p) \leq \frac{F''(x)}{(1-\alpha\,\sigma)^2}.$$

En faisant tendre σ vers 0, autrement dit vers $\sigma_x(p)$ et par passage à la limite dans les inégalités, nous aboutissons à l'inégalité escomptée, à savoir :

$$F''(x - \alpha\,p) \leq \frac{F''(x)}{(1-\alpha\,\sigma_x(p))^2}.$$

3. Nous démontrons la première inégalité de (3.12) dans le cas où $\sigma_x(-p)$ est nul. Etant donné v un élément de K, l'application de la seconde inégalité de (3.12) avec l'identification $p = -v$ aboutit à l'inégalité suivante :

$$\forall\,\alpha \in [0, \frac{1}{\sigma_x(-v)}) \qquad F''(x + \alpha\,v) \leq \frac{F''(x)}{(1-\alpha\,\sigma_x(-v))^2}$$

qui, par les relations (3.5) et (3.3), se simplifie en

$$\forall\,\alpha \in [0, +\infty) \qquad F''(x + \alpha\,v) \leq F''(x).$$

La particularisation de cette inégalité au cas $\alpha = 1$ nous conduit au résultat suivant :

$$F''(x + v) \leq F''(x).$$

Dès lors, la première inégalité de (3.12), dans le cas où $\sigma_x(-p)$ est nul, est obtenue par application de ce résultat avec les identifications $x = x - \alpha\,p$ et $v = \alpha\,p$. Remarquons que l'appartenance de $\alpha\,p$ à K est garantie par la relation (3.5).

4. Nous démontrons la première inégalité de (3.12) dans le cas où $\sigma_x(-p)$ est strictement positif. Nous montrons, tout d'abord, l'assertion suivante :

$$\sigma_{x-\alpha p}(-p) \leq \frac{\sigma_x(-p)}{(1+\alpha\,\sigma_x(-p))}.$$

Pour ce faire, il suffit de montrer l'appartenance ci-dessous :

$$\frac{\sigma_x(-p)}{(1+\alpha\,\sigma_x(-p))} \in \{\beta \geq 0 : [\beta\,(x-\alpha\,p)] + p \in K\}.$$

La positivité du rapport découle de celles de $\sigma_x(-p)$ et de α. D'autre part, nous vérifions la satisfaction de la relation ci-dessous :

$$\frac{\sigma_x(-p)}{(1+\alpha\,\sigma_x(-p))}\,(x-\alpha\,p)+p \quad \in K.$$

Remarquons que

$$\begin{aligned}
&\frac{\sigma_x(-p)}{(1+\alpha\,\sigma_x(-p))}\,(x-\alpha\,p)+p \\
&\quad = \quad (\frac{\sigma_x(-p)}{1+\alpha\,\sigma_x(-p)})\,[(x-\alpha\,p)+\frac{1+\alpha\,\sigma_x(-p)}{\sigma_x(-p)}\,p] \\
&\quad = \quad (\frac{\sigma_x(-p)}{1+\alpha\,\sigma_x(-p)})\,[(x-\alpha\,p)+(\alpha+\frac{1}{\sigma_x(-p)})\,p].
\end{aligned}$$

Par conséquent, nous pouvons nous contenter de montrer que $(x-\alpha\,p)+(\alpha+\frac{1}{\sigma_x(-p)})\,p$ appartient à K. Or, cette quantité peut se réécrire comme $x-\frac{1}{\sigma_x(-p)}\,(-p)$ qui, par définition de $\sigma_x(-p)$, est bien un élément de K.
Nous montrons, ensuite, l'inégalité suivante :

$$F''(x) \leq \frac{F''(x-\alpha\,p)}{(1-\alpha\sigma_{x-\alpha\,p}(-p))^2}.$$

Dans ce but, nous appliquons la seconde inégalité de (3.12) avec l'identification $p=-p$. De la sorte, nous obtenons, pour tout α appartenant à $[0, \frac{1}{\sigma_x(-p)})$, l'inégalité ci-dessous :

$$F''(x+\alpha\,p) \leq \frac{F''(x)}{(1-\alpha\,\sigma_x(-p))^2}.$$

Nous effectuons alors le changement de variables $y = x+\alpha\,p$ et aboutissons à la relation voulue, à savoir :

$$F''(y) \leq \frac{F''(y-\alpha\,p)}{(1-\alpha\sigma_{y-\alpha\,p}(-p))^2}.$$

Nous nous attachons finalement à expliciter l'obtention de la première inégalité de

(3.12) dans le cas où $\sigma_x(-p)$ est strictement positif. Puisque nous disposons de l'inégalité suivante :

$$(1-\alpha\sigma_{x-\alpha p}(-p))^2\; F''(x) \le F''(x-\alpha\, p),$$

il est suffisant de montrer que

$$\frac{1}{(1+\alpha\,\sigma_x(-p))^2} \le (1-\alpha\,\sigma_{x-\alpha\, p}(-p))^2.$$

Or, dans un premier temps, nous avons obtenu la relation ci-dessous :

$$\sigma_{x-\alpha\, p}(-p) \le \frac{\sigma_x(-p)}{(1+\alpha\,\sigma_x(-p))},$$

qui conduit, étant donné la stricte positivité de α, à l'inégalité suivante :

$$1-\alpha\,\sigma_{x-\alpha\, p}(-p) \ge \frac{1}{1+\alpha\,\sigma_x(-p)}.$$

Nous aboutissons, dès lors, au résultat escompté en élevant au carré les deux membres de cette inégalité puisque $\frac{1}{1+\alpha\,\sigma_x(-p)}$ est strictement positif.

$\square$

Nous rappelons l'analogue du Théorème 3.1 établi par Yu. Nesterov et A. Nemirovskii dans [2, Théorème 2.1.1] pour des barrières seulement self-concordantes :

Théorème 2 *Quels que soient x un point intérieur de K fixé et p un élément de E, nous avons, pour tout ω point intérieur de K situé à l'intérieur de la boule unité centrée en x et générée par la norme locale définie par le hessien de F en x, la double inégalité suivante :*

$$(1-r^2)\; D^2F(x)\,[p,p] \le D^2F(\omega)\,[p,p] \le \frac{1}{1-r^2}\; D^2F(x)\,[p,p]$$

où $r \overset{def}{=} \|x-\omega\|_x$.

Souvenons-nous que Yu. Nesterov et A. Nemirovskii évoluent dans un cadre de travail beaucoup plus général que le nôtre : d'une part, la seule mesure de la distance d'un point de l'espace à la frontière de K relativement à un point intérieur de celui-ci qu'ils utilisent est la norme locale définie par le hessien de la barrière en ce point; d'autre part, ils considèrent des barrières seulement self-concordantes pour K. Dans ces conditions, le hessien de F en un point ω ne peut être borné, en termes de hessien de F

en x, que si ω se situe à l'intérieur de la boule unité centrée en x et définie par la norme locale.

En nous restreignant aux barrières "ν-self-scaled" et en définissant de nouvelles mesures de la distance d'un point de l'espace à la frontière de K relativement à un point intérieur de celui-ci, nous avons pu affiner considérablement ce résultat. Effectivement, le Théorème 3.1 nous montre que l'évolution dans ce cadre de travail permet de borner le hessien de F, en termes de hessien de F en x, en tout point de l'intérieur de K.

Corollaire 3.1

(i.) Nous avons, quels que soient x un point intérieur de K, v un élément de K et α un réel appartenant à [0,1), la double inégalité suivante :

$$F''(x+v) \leq F''(x) \leq F''(x - \frac{\alpha}{\sigma_x(v)} v).$$

(ii.) Nous avons, quels que soient x et ω des points intérieurs de K, l'inégalité ci-dessous :

$$F''(x) \leq \sigma_x^2(\omega)\ F''(\omega).$$

<u>Preuve:</u>

(i.) Remarquons, tout d'abord, que la première inégalité à démontrer a d'ores et déjà été obtenue dans le troisième point de la preuve du Théorème 3.1. Il nous reste donc à prouver la seconde inégalité. Nous définissons z par

$$z \stackrel{def}{=} x - \frac{\alpha}{\sigma_x(v)} v$$

et nous appliquons la première inégalité de (3.12) avec l'identification $p = x - z$. De la sorte, nous obtenons l'inégalité suivante :

$$\forall\, \alpha \in [0, \frac{1}{\sigma_x(x-z)}) \qquad \frac{1}{(1+\alpha\,\sigma_x(-(x-z)))^2} F''(x) \leq F''(x - \alpha\,(x-z)).$$

Comme $x - z$ appartient à K, l'équivalence (3.5) implique l'annulation de $\sigma_x(-(x-z))$. Par ailleurs, nous avons le résultat ci-dessous :

$$\frac{1}{\sigma_x(x-z)} \geq 1.$$

Celui-ci découle de la relation (3.4) et de l'appartenance de z à K. Par conséquent, tenant compte des deux dernières constatations mentionnées, nous obtenons l'inégalité suivante :

$$\forall\, \alpha \in [0,1) \qquad F''(x) \leq F''(x - \alpha\,(x - z)).$$

Nous faisons alors tendre α vers 1. Puisque F est entre autres une barrière ν-logarithmiquement homogène pour K, elle est de classe C^2 et donc, par passage à la limite dans les inégalités, nous aboutissons à la relation ci-dessous :

$$F''(x) \leq F''(z)$$

qui conduit bien au résultat escompté, étant donnée la définition de z.

($ii.$) Nous définissons v un élément de K par

$$v \stackrel{def}{=} \sigma_x(\omega)\, x - \omega$$

et nous appliquons la première inégalité de ($i.$) avec l'identification $x = \omega$. De la sorte, nous obtenons le résultat suivant :

$$F''(\omega + v) \leq F''(\omega)$$

qui, par définition de v, est équivalent à l'inégalité ci-dessous :

$$F''(\sigma_x(\omega)x) \leq F''(\omega).$$

Par ailleurs, la relation (2.2) implique l'égalité suivante :

$$F''(\sigma_x(\omega)\, x) = \frac{1}{(\sigma_x(\omega))^2} F''(x).$$

Par conséquent, la combinaison des deux dernières relations obtenues conduit à l'inégalité suivante :

$$\frac{1}{(\sigma_x(\omega))^2} F''(x) \leq F''(\omega),$$

qui peut être réécrite de façon équivalente comme suit :

$$F''(x) \leq (\sigma_x(\omega))^2 F''(\omega).$$

□

3.2.2 Borne sur la variation d'une barrière "ν-self-scaled"

Nous nous intéressons, tout d'abord, au cas où $\sigma_x(p)$ est strictement positif.

Théorème 3.2 *Quels que soient x un point intérieur de K fixé et p un élément de E tel que $\sigma_x(p)$ est strictement positif, nous avons, pour tout α appartenant à $[0, \frac{1}{\sigma_x(p)})$, l'inégalité suivante :*

$$F(x - \alpha p) \quad \leq \quad F(x) - \alpha < F'(x), p > + \frac{\|p\|_x^2}{\sigma_x^2(p)}(-\alpha\, \sigma_x(p) - \ln[1 - \alpha\, \sigma_x(p)]) \quad (3.13)$$

__Preuve:__
Nous définissons $\bar{p}$ comme décrit dans l'énoncé du Lemme 3.1, à savoir

$$\bar{p} \stackrel{def}{=} \frac{-p}{\sigma_x(p)}$$

et nous considérons la fonction définie sur l'intervalle $[0, 1)$ par

$$\theta(\beta) \stackrel{def}{=} F(x(\beta)),$$

où $x(\beta) \stackrel{def}{=} x + \beta\, \bar{p}$.
Nous démontrons tout d'abord, pour tout β appartenant à $[0, 1)$, l'inégalité suivante :

$$\theta''(\beta) \leq \frac{\theta''(0)}{(1-\beta)^2}$$

Nous obtenons, par application du Lemme 3.1 et par explicitation de $\phi(\beta)$, l'inégalité suivante :

$$< F''(x(\beta))\; v, v > \leq \frac{< F''(x(0))\; v, v >}{(1-\beta)^2},$$

où v est une direction de E arbitrairement fixée. Puisque $\bar{p}$ est un élément de E, nous pouvons particulariser cette inégalité à $v = \bar{p}$ et obtenir de la sorte la relation ci-dessous :

$$< F''(x(\beta))\; \bar{p}, \bar{p} > \leq \frac{< F''(x(0))\; \bar{p}, \bar{p} >}{(1-\beta)^2}.$$

Cette dernière conclut bien au résultat intermédiaire escompté puisque la dérivée seconde de la fonction θ vaut l'expression suivante :

$$< F''(x(\beta))\; \bar{p}, \bar{p} > .$$

Nous prouvons alors l'inégalité (3.13). Par intégrations successives et emploi du résultat intermédiaire obtenu en premier lieu, nous obtenons le développement suivant :

$$
\begin{aligned}
\theta(\beta) - \theta(0) &= \int_0^\beta \theta'(\lambda)\, \mathrm{d}\lambda \\
&= \theta'(0)\, \beta + \int_0^\beta \int_0^\lambda \theta''(\tau)\, \mathrm{d}\tau\, \mathrm{d}\lambda \\
&\leq \theta'(0)\, \beta + \int_0^\beta \int_0^\lambda \frac{\theta''(0)}{(1-\tau)^2}\, \mathrm{d}\tau\, \mathrm{d}\lambda \\
&= \theta'(0)\, \beta + \theta''(0)(-\beta - \ln(1-\beta)).
\end{aligned}
$$

Nous particularisons celui-ci à $\beta = \alpha\, \sigma_x(p)$. La validité de cette affectation est garantie par l'hypothèse d'appartenance de α à $[0, \frac{1}{\sigma_x(p)})$. L'inégalité finale de ce développement devient, alors, en explicitant $\theta(\beta)$, $x(\beta)$ et $\bar{p}$, la suivante :

$$
\begin{aligned}
F(x - \alpha\, p) - F(x) \;\leq\; & -\alpha\, \sigma_x(p) < F'(x), \frac{p}{\sigma_x(p)} > \\
& + < F''(x)\frac{p}{\sigma_x(p)}, \frac{p}{\sigma_x(p)} > (-\alpha\, \sigma_x(p) - \ln[1 - \alpha\, \sigma_x(p)]).
\end{aligned}
$$

Par la linéarité du produit scalaire et la définition de la norme $\|.\|_x$, nous aboutissons à l'inégalité voulue.

$\square$

Proposition 3.2

(i) $\forall\, \tau \geq 0$ $\quad \tau - \ln(1+\tau)$ *est monotonement croissante*

(ii) $\forall\, \tau > -1$ $\quad \frac{\tau - \ln(1+\tau)}{\tau^2}$ *(définie comme* $\frac{1}{2}$ *si* $\tau = 0$*) est monotonement décroissante*

(iii) $\forall\, \tau < 1$ $\quad \frac{-\tau - \ln(1-\tau)}{\tau^2}$ *(définie comme* $\frac{1}{2}$ *si* $\tau = 0$*) est monotonement croissante.*

Le Théorème 3.2 fournit effectivement une borne sur la variation d'une barrière "ν-self-scaled". Celle-ci, couplée à la Proposition 3.2, permettra la démonstration de théorèmes qui contribuent à garantir la convergence des méthodes de réduction de potentiel primales et primales-duales. Nous remarquons que l'introduction des nouvelles distances à la frontière de K et l'emploi de cônes et de barrières "self-scaled" valident la borne fournie pour une longueur de pas α dans la direction p supérieure à celle qui aurait été autorisée lorsque seule la norme locale définie par le hessien de la barrière et une barrière self-concordante étaient disponibles. Cette supériorité laisse présager d'une augmentation possible de la longueur de pas dans les algorithmes utilisant des barrières "ν-self-scaled" plutôt que des barrières self-concordantes.

Dans le cas où $\sigma_x(p)$ est nul, on peut montrer que, pour tout α appartenant à $[0, +\infty)$, on obtient l'inégalité ci-dessous :

$$F(x + \alpha\, p) \leq F(x) + \alpha < F'(x), p > + \frac{\alpha^2}{2} \|p\|_x^2.$$

Chapitre 4

Comportement d'une barrière "ν-self-scaled" sur un cône bidimensionnel défini par des directions orthogonales

Nous définissons, tout d'abord, les notions de directions orthogonales et de cône bidimensionnel défini par de telles directions. De plus, nous décrivons le comportement d'une barrière "ν-self-scaled" sur un tel cône.

Cette étude nous permet alors d'établir un résultat dont nous pourrons apprécier la portée lors de l'élaboration d'une méthode de réduction de potentiel primale-duale.

4.1 Examen du comportement d'une barrière "ν-self-scaled" sur un cône bidimensionnel défini par des directions orthogonales

Définition 4.1 *Soient ω un point intérieur de K fixé ainsi que v et z des éléments de la frontière de K. z est dit être une direction orthogonale de v relativement à ω si et seulement si*

$$< F''(\omega)v, z > \; = \; 0.$$

D'une part, on peut montrer, grâce au point (*iii.*) du Théorème 2.1, que l'emploi de cônes et de barrières "self-scaled" suffit à garantir l'existence de telles directions. D'autre part, l'hypothèse de symétrie imposée aux hessiens de la barrière "ν-self-scaled" F considérée, implique l'équivalence suivante : z est orthogonal à v relativement à ω si et seulement si v est orthogonal à z relativement à ω.

Définition 4.2 *Nous considérons ω un point intérieur de K fixé ainsi que v et z deux éléments de la frontière de K tels que v est orthogonal à z relativement à ω. Nous définissons D le cône bidimensionnel défini par v et z comme suit*

$$D \stackrel{def}{=} \{x : x = \omega + \alpha v + \beta z, \quad \alpha \geq 0 \text{ et } \beta \geq 0\}.$$

Puisque nous nous intéressons au comportement de la barrière "ν-self-scaled" F sur le cône D, nous devons, avant toute chose, vérifier que F est bien définie sur D. Autrement dit, D doit être contenu dans l'intérieur de K, le domaine de définition de F. L'appartenance à l'intérieur de K de $\omega + \alpha v + \beta z$, quels que soient α et β positifs, est obtenue en deux étapes. D'une part, les relations (3.3) et (3.5) permettent de démontrer que $\omega + \alpha v$ est un point intérieur de K quel que soit α positif. D'autre part, on fixe un α positif, on pose $\bar{\omega} = \omega + \alpha v$ et l'on a à nouveau recours aux relations (3.3) et (3.5) pour montrer l'appartenance de $\bar{\omega} + \beta z$ à l'intérieur de K pour tout β positif.

Nous pouvons, désormais, étudier le comportement de F sur D.

Lemme 4.1 *Quel que soit x un élément du cône D, nous avons les égalités suivantes :*

$$< F''(x)v, z > \; = \; 0 \tag{4.1}$$
$$< F'''(x)[z]v, v > \; = \; 0 \tag{4.2}$$
$$< F'''(x)[v]z, z > \; = \; 0.$$

<u>Preuve:</u>

(4.1) Dans un premier temps, nous démontrons, pour tout α positif, l'assertion suivante :

$$< F''(\omega + \alpha v)v, z > \; = \; 0.$$

Dans ce but, nous considérons la fonction ψ définie sur $I\!R^+$ par

$$\psi(\alpha) \stackrel{def}{=} - < F'(\omega + \alpha v), z >$$

et nous démontrons sa convexité par application du Lemme 2.1 avec les identifications $x = \omega + \alpha v$ et $v = z$. Dès lors, nous obtenons, quel que soit α positif, l'inégalité suivante :

$$\psi'(\alpha) \geq \psi'(0).$$

La dérivée première de ψ évaluée en un α positif quelconque vaut l'expression suivante :

$$- < F''(\omega + \alpha v)v, z > .$$

Dès lors, l'orthogonalité de z par rapport à v relativement à ω implique l'annulation de cette dérivée lorsqu'elle est évaluée en zéro. L'inégalité déduite de la convexité de ψ et vérifiée par tout α positif, devient, dès lors, la suivante :

$$< F''(\omega + \alpha v)v, z > \leq 0.$$

Par ailleurs, le point (*iii.*) du Théorème 2.1, couplé à la définition de K^*, implique la satisfaction par tout α positif de la relation ci-dessous :

$$< F''(\omega + \alpha v)v, z > \geq 0.$$

La combinaison des deux dernières inégalités obtenues conduit alors à l'assertion intermédiaire escomptée.
Ensuite, nous fixons un α positif, nous posons $\bar{\omega} = \omega + \alpha v$ et nous démontrons, quel que soit β positif, l'égalité ci-dessous :

$$< F''(\bar{\omega} + \beta z)v, z > = 0.$$

Pour ce faire, nous considérons une nouvelle fonction ψ définie sur $I\!R^+$ par

$$\psi(\beta) \stackrel{def}{=} - < F'(\bar{\omega} + \beta z), v > .$$

La convexité de celle-ci découle, à nouveau, de l'application du Lemme 2.1 mais, cette fois, avec les identifications $x = \bar{\omega} + \beta z$ et $v = v$. Cette caractéristique de la fonction ψ entraîne la satisfaction par tout β positif de l'inégalité suivante :

$$\psi'(\beta) \geq \psi'(0).$$

De nouveau, l'évaluation de la dérivée première de ψ en zéro est nulle. Ce fait s'explique par l'hypothèse de symétrie imposée aux hessiens de F et par le résultat démontré en premier lieu. L'inégalité déduite de la convexité de ψ et vérifiée par tout β positif, devient, dès lors, la suivante :

$$< F''(\bar{\omega} + \beta z)z, v > \leq 0.$$

Par ailleurs, le point (*iii.*) du Théorème 2.1 et la définition de K^* impliquent la vérification par tout β positif de la relation ci-dessous :

$$< F''(\bar{\omega} + \beta z)z, v > \geq 0.$$

La combinaison des deux dernières inégalités obtenues conduit alors au résultat escompté.
De la sorte, nous avons montré l'égalité suivante :

$$< F''(\omega + \alpha v + \beta z)v, z >= 0$$

quels que soient α et β positifs et donc, la propriété (4.1) quel que soit x appartenant à D.

(4.2) Nous démontrons, tout d'abord, la seconde égalité. Celle-ci est obtenue en dérivant l'égalité (4.1) dans la direction z. En effet, le premier membre de l'égalité décrite en (4.1) coïncide avec la dérivée directionnelle d'ordre deux de F en x dans les directions v et z.

La première inégalité est, quant à elle, obtenue en dérivant dans la direction v l'égalité ci-dessous :

$$< F''(x)z, v >= 0,$$

dont le premier membre vaut la dérivée directionnelle d'ordre deux de F en x dans les directions z et v. Remarquons que la propriété (4.1) et l'hypothèse de symétrie imposée aux hessiens de F garantissent la vérification de l'égalité ci-dessus.

□

La relation (4.1) montre que les directions v et z, qui étaient supposées orthogonales relativement à ω, le sont en fait relativement à tout point du cône D qu'elles définissent. Nous utilisons maintenant ce résultat pour décrire le comportement de F sur D.

Théorème 4.1 *Quels que soient α et β positifs, les égalités suivantes sont satisfaites par l'élément x de D défini par $x = \omega + \alpha v + \beta z$:*

$$\begin{aligned} F(x) &= F(\omega + \alpha v) + F(\omega + \beta z) - F(\omega) && (4.3) \\ < F'(x), v > &= < F'(\omega + \alpha v), v > && (4.4) \\ < F'(x), z > &= < F'(\omega + \beta z), z > && \\ < F''(x)v, v > &= < F'(\omega + \alpha v)v, v > && \\ < F''(x)v, z > &= 0 && \\ < F''(x)z, z > &= < F''(\omega + \beta z)z, z > && (4.5) \end{aligned}$$

<u>Preuve:</u>

(4.3) Des intégrations successives et l'utilisation de l'égalité (4.1) justifient le développement suivant :

$$\begin{aligned} & F(x) - F(\omega + \alpha v) \\ &= \int_0^\beta < F'(\omega + \alpha v + \tau z), z > \; \mathrm{d}\tau \\ &= \int_0^\beta [< F'(\omega + \tau z), z > + \int_0^\alpha < F''(\omega + \rho v + \tau z)v, z > \; \mathrm{d}\rho] \; \mathrm{d}\tau \\ &= \int_0^\beta < F'(\omega + \tau z), z > \; \mathrm{d}\tau \\ &= F(\omega + \beta z) - F(\omega). \end{aligned}$$

qui conclut bien à l'égalité décrite en (4.3).

(4.4) Les quatre égalités à démontrer s'obtiennent par dérivation de l'assertion (4.3) par rapport aux arguments adéquats. Les dérivées premières respectivement par rapport à α et β fournissent les deux premières égalités. La troisième s'obtient en dérivant deux fois par rapport à α. Quant à la quatrième, elle résulte d'une dérivation de l'égalité décrite en (4.3) par rapport à β suivie d'une autre par rapport à α.

(4.5) Cette égalité est obtenue en dérivant deux fois la propriété (4.3) par rapport à β.

□

Le Théorème 4.1 montre que la barrière "ν-self-scaled" F peut être séparée sur le plan défini par les directions orthogonales v et z, et passant par ω. Le Théorème 4.2 exploite cette propriété afin de dégager une inégalité d'intérêt considérable dans les méthodes de réduction de potentiel primales-duales.

4.2 Exploitation du comportement particulier d'une barrière "ν-self-scaled" sur un cône bidimensionnel défini par des directions orthogonales

Théorème 4.2 *Quels que soient x et s des points intérieurs respectifs de K et de K^*, nous avons l'inégalité suivante :*

$$< F'(x), F_*'(s) > \;\; \geq \;\; \frac{\nu\,(\nu-1)}{< s, x >} + \frac{3}{4}\,\sigma_x^2(\omega) \tag{4.6}$$

où ω est un point intérieur de K tel que $s = F''(\omega)\,x$.

<u>Preuve:</u>

1. Nous prouvons qu'il est suffisant de montrer, pour tout x et s points intérieurs respectifs de K et de K^*, l'inégalité suivante :

$$(\|F'(\omega+v)\|_\omega^*)^2 \;\; \geq \;\; \frac{(\nu-1)^2}{\|\omega+v\|_\omega^2 - 1} + 1 \tag{4.7}$$

où ω est un point intérieur de K tel que $s = F''(\omega)\,x$ et où v est un élément de la frontière de K défini par $v \stackrel{def}{=} \sigma_x(\omega)\,x - \omega$. L'existence d'un point intérieur ω de

K tel que $s = F''(\omega)\, x$, est garantie par le Théorème 2.2 qui implique également la satisfaction de l'assertion suivante :

$$F'(x) = F''(\omega)\, F_*'(s).$$

Sur base de ces deux égalités et de la définition des normes $\|.\|_\omega^*$ et $\|.\|_\omega$, nous pouvons réécrire l'inégalité (4.6) du Théorème 4.2 comme suit :

$$(\|F'(x)\|_\omega^*)^2 \quad \geq \quad \frac{\nu\,(\nu-1)}{\|x\|_\omega^2} + \frac{3}{4}\,\sigma_x^2(\omega). \tag{4.8}$$

D'autre part, la définition de v entraîne l'identité suivante :

$$x = \frac{\omega + v}{\sigma_x(\omega)}.$$

Par conséquent, la relation (4.8) se transforme, grâce également à la première égalité de (1.1), en l'inégalité suivante :

$$(\|F'(\omega+v)\|_\omega^*)^2 \quad \geq \quad \frac{\nu\,(\nu-1)}{\|\omega+v\|_\omega^2} + \frac{3}{4}.$$

Dès lors, montrer l'assertion (4.7) suffit à prouver le Théorème 4.2 si la condition suivante :

$$\frac{(\nu-1)^2}{\|\omega+v\|_\omega^2 - 1} + 1 \geq \frac{\nu\,(\nu-1)}{\|\omega+v\|_\omega^2} + \frac{3}{4}$$

est vérifiée. Nous montrons donc la satisfaction de celle-ci. Dans ce but, nous posons $\rho \stackrel{def}{=} \|\omega+v\|_\omega^2$. Nous appliquons la première partie du point ($i.$) du Corollaire 3.1 avec l'identification $x = \omega$. Nous obtenons de la sorte l'inégalité suivante :

$$F''(\omega+v) \leq F''(\omega)$$

qui, par définition des inégalités relatives au cône des opérateurs semi-définis positifs, implique la relation ci-dessous :

$$< F''(\omega+v)(\omega+v), \omega+v > \;\leq\; < F''(\omega)(\omega+v), \omega+v > .$$

Nous obtenons alors, par définition de ρ et de la norme $\|.\|_\omega$, l'inégalité suivante :

$$\rho \;\geq\; < F''(\omega+v)(\omega+v), \omega+v >$$

dont le second membre vaut ν par la première égalité décrite en (1.4). Rappelons que ν est un réel supérieur ou égal à 1. Le développement ci-dessous transforme la condition à satisfaire en une inégalité plus aisément vérifiable. Nous y employons la double inégalité $\rho \geq \nu \geq 1$ obtenue ci-dessus pour justifier la conservation du sens de l'inégalité lors de l'élimination du dénominateur commun et pour vérifier la satisfaction de la dernière inégalité. Cette vérification clôture le premier pas de la preuve du Théorème 4.2.

$$
\begin{array}{rcl}
\dfrac{(\nu-1)^2}{\rho-1}+1 & \geq & \dfrac{\nu(\nu-1)}{\rho}+\dfrac{3}{4} \\
& \Updownarrow & \\
\dfrac{(\nu-1)^2}{\rho-1}+\dfrac{1}{4}-\dfrac{\nu(\nu-1)}{\rho} & \geq & 0 \\
& \Updownarrow & \\
\rho(\nu-1)^2+\dfrac{1}{4}\rho(\rho-1)-\nu(\nu-1)(\rho-1) & \geq & 0 \\
& \Updownarrow & \\
\dfrac{1}{4}\rho(\rho-1)-(\nu-1)(\rho-\nu) & \geq & 0 \\
& \Updownarrow & \\
\dfrac{1}{4}(\rho-1+[(\rho-\nu)-(\nu-1)]^2) & \geq & 0.
\end{array}
$$

2. Nous démontrons l'inégalité (4.7). Dans ce but, nous considérons z une direction orthogonale à v relativement à ω et nous définissons les deux quantités suivantes :

$$
\begin{array}{rcl}
\mu & \stackrel{def}{=} & \dfrac{\rho-\nu}{\nu-1}, \\
\tau & \stackrel{def}{=} & \dfrac{-\mu}{< F'(\omega+v), z >}.
\end{array}
$$

Nous nous intéressons alors à l'identité ci-dessous :

$$< F'(\omega+v), \omega+v+\tau z > \; = \; < F'(\omega+v), \omega+v > +\tau < F'(\omega+v), z > .$$

L'égalité décrite en (1.3) et la particularisation à $\alpha = 1$ et $\beta = 0$ de la seconde égalité répertoriée en (4.4) permettent respectivement une réécriture du premier et du second produit scalaire intervenant dans le deuxième membre de l'identité considérée. Ces manipulations transforment cette identité en l'égalité ci-dessous :

$$< F'(\omega+v), \omega+v+\tau z > \; = \; -\nu+\tau < F'(\omega), z >$$

qui, par définition de τ, devient la suivante :

$$< F'(\omega+v), \omega+v+\tau z > \; = \; -\nu-\mu.$$

Par conséquent, en élevant au carré les deux membres de cette égalité et en appliquant l'inégalité de Cauchy-Scharwz généralisée aux espaces duaux, nous obtenons la relation suivante :

$$(\nu+\mu)^2 \;\leq\; (\|F'(\omega+v)\|_\omega^*)^2 \; \|\omega+v+\tau z\|_\omega^2$$

et, par suite, l'inégalité ci-dessous :

$$(\|F'(\omega+v)\|_\omega^*)^2 \;\geq\; \frac{(\nu+\mu)^2}{\|\omega+v+\tau z\|_\omega^2}. \tag{4.9}$$

Dans un premier temps, nous majorons le dénominateur intervenant dans (4.9). La linéarité du produit scalaire, la définition de ρ et l'orthogonalité de z par rapport à v relativement à ω conduisent à l'égalité suivante :

$$\|\omega+v+\tau z\|_\omega^2 = \rho + 2\tau < F''(\omega)\,\omega, z > +\tau^2\|z\|_\omega^2.$$

L'explicitation de τ et de la norme $\|.\|_\omega$ ainsi que l'utilisation de la première égalité décrite en (1.2) permettent la réécriture suivante du second membre de cette égalité :

$$\rho + 2\mu + \mu^2 \frac{< F''(\omega)\,z, z >}{< F'(\omega), z >^2}.$$

Par ailleurs, nous montrons l'inégalité suivante :

$$\frac{< F''(\omega)z, z >}{< F'(\omega), z >^2} \leq 1.$$

Pour ce faire, nous rappelons un résultat obtenu par Yu. Nesterov et A. Nemirovskii dans [2,Corollaire 2.3.1].

Corollaire 1 *Soient ν un réel positif, C une partie convexe, fermée, à intérieur non vide de E et F une barrière ν-self-concordante pour C. Nous considérons également p une direction de E telle que x+tp appartient à C, quels que soient x un point intérieur de C et t un réel strictement positif.*
Alors, nous avons, pour tout x point intérieur de C, l'inégalité suivante :

$$(D^2F(x)[p,p])^{\frac{1}{2}} \leq -DF(x)[p].$$

Nous appliquons ce résultat au cône K et à la fonction F considérés ainsi qu'à la direction z. Nous avons déjà souligné, grâce aux relations (3.3) et (3.5), l'appartenance à l'intérieur de K de tout élément de la forme $\omega + tz$ où ω est un point intérieur de K et t un réel strictement positif. Pour satisfaire complètement les hypothèses du Corollaire 1, nous devons étendre ce résultat à K, c'est-à-dire montrer que tout élément de la forme $\omega + tz$ où ω est un élément de K et t un réel strictement positif, appartient à K. Cette appartenance est obtenue par extention à la frontière de K de la définition de $\sigma_x(p)$ et de la relation (3.3) ainsi que par la relation (3.5). L'application du Corollaire 1 conclut alors à la satisfaction par tout ω point intérieur de K de l'inégalité ci-dessous :

$$< F''(\omega)\,z, z >^{\frac{1}{2}} \;\leq\; - < F'(\omega), z >$$

qui conduit au résultat escompté par élévation au carré. Par conséquent, nous obtenons la majoration suivante du dénominateur intervenant dans (4.9) :

$$\|\omega + v + \tau z\|_\omega^2 \;\leq\; \rho + 2\mu + \mu^2.$$

La définition de μ permet de réécrire la borne de cette majoration comme suit :

$$\begin{aligned}
\rho + 2\mu + \mu^2 &= \rho - 1 + (\mu + 1)^2 \\
&= (\rho - 1)(1 + \frac{\rho - 1}{(\nu - 1)^2}) \\
&= \frac{(\rho - 1)(\rho - 2\nu + \nu^2)}{(\nu - 1)^2},
\end{aligned}$$

ce qui conduit à la majoration finale suivante :

$$\|\omega + v + \tau z\|_\omega^2 \;\leq\; \frac{(\rho - 1)(\rho - 2\nu + \nu^2)}{(\nu - 1)^2}.$$

Par ailleurs, une réécriture du numérateur intervenant dans (4.9) est rendue possible à nouveau par la définition de μ. En effet, on vérifie facilement l'égalité ci-dessous :

$$\nu + \mu = \frac{\rho - 2\nu + \nu^2}{\nu - 1}.$$

La majoration et la réécriture respectivement du dénominateur et du numérateur intervenant dans l'inégalité (4.9) permettent alors de conclure à la minoration attendue du premier membre de cette inégalité de la façon suivante :

$$\begin{aligned}
(\|F'(\omega + v)\|_\omega^*)^2 &\geq \frac{(\nu + \mu)^2}{\|\omega + v + \tau z\|_\omega^2} \\
&\geq \frac{(\rho - 2\nu + \nu^2)^2(\nu - 1)^2}{(\nu - 1)^2(\rho - 1)(\rho - 2\nu + \nu^2)} \\
&= \frac{\rho - 2\nu + \nu^2}{\rho - 1} \\
&= 1 + \frac{1 - 2\nu + \nu^2}{\rho - 1} \\
&= 1 + \frac{(\nu - 1)^2}{\rho - 1}.
\end{aligned}$$

□

L'inégalité (4.6) décrite dans le Théorème 4.2 permettra la démonstration d'un théorème qui contribue à garantir la convergence à grands pas de la méthode de réduction de potentiel primale-duale. Par conséquent, nous constatons que l'étude d'une propriété particulière des barrières "ν-self-scaled" justifie, par la portée du résultat qu'elle permet de déduire, la considération de telles barrières dans les méthodes de réduction de potentiel primales-duales.

Dans le chapitre suivant, nous achevons la définition de notre cadre de travail et étoffons encore de quelques propositions notre éventail de résultats utiles pour la suite.

Chapitre 5

Spécification des problèmes considérés et étude des projections obliques

La première partie de ce chapitre précise le type de problèmes primal et dual que nous considérerons dans la suite et répertorie nos hypothèses ainsi que leurs conséquences directes.

Nous étudions ensuite les projections obliques. Cette étude comprend la définition et surtout l'établissement des propriétés et d'une méthode de calcul de ces projections. Notre intérêt pour celles-ci s'explique par le fait que les systèmes linéaires les définissant coïncident avec ceux que nous aurons à résoudre dans nos algorithmes lors de la détermination des directions de recherche.

5.1 Définition du cadre de travail

Le problème primal (P) considéré est le suivant :

$$\left.\begin{array}{l} \min < c, x > \\ sc. \\ \quad Ax = b \\ \quad x \in K \end{array}\right\} \quad (P)$$

où A désigne un opérateur linéaire surjectif défini sur E et à valeurs dans un espace vectoriel de dimension finie noté Y^*. Dès lors, c est un élément de E^* et b un élément de Y^*. Le problème dual (D) associé à (P) est le suivant (voir [7]) :

$$\left.\begin{array}{l} \max < b, y > \\ sc. \\ \quad A^*y + s = c \\ \quad s \in K^* \end{array}\right\} \quad (D)$$

où A^* désigne l'adjointe de A et est donc une application linéaire injective définie sur Y et à valeurs dans E^*. Dès lors, y est un élément de Y.

Les hypothèses de linéarité et de surjectivité imposées à l'opérateur A justifient le développement d'algorithmes uniquement primal et primal-dual et établissent la suffisance de la détermination d'une variable d'écart s optimale lors de la résolution du dual. En effet, sous ces hypothèses, on peut montrer que le problème dual (D) se ramène à un problème de la forme (P) uniquement en les variables d'écart s et également que, pour tout s élément de K^*, il existe un et un seul y appartenant à Y tel que l'égalité suivante est satisfaite :

$$A^*y + s = c.$$

Par ailleurs, soulignons le caractère non restrictif de l'hypothèse de surjectivité. Dans le cas où celle-ci n'est pas satisfaite, nous restreignons simplement Y^* à l'image de A. On peut montrer que cette restriction ne modifie pas les ensembles de solutions de (P) et de (D).

Nous supposons l'existence d'un point strictement admissible pour le primal et pour le dual. Nous avons dès lors que les deux ensembles définis par

$$\begin{array}{l} S^0(P) \stackrel{def}{=} \{x \in \text{int } K : \ Ax = b\} \\ S^0(D) \stackrel{def}{=} \{(y, s) \in Y \times \text{int } K^* : \ A^*y + s = c\} \end{array}$$

sont non vides. Cette double hypothèse nous assure, d'une part, que les problèmes primal et dual (P) et (D) admettent des solutions et que leurs valeurs optimales coïncident. D'autre part, elle implique le caractère borné des ensembles de solutions de ces deux problèmes. Le théorème de dualité établi par Yu. Nesterov et A. Nemirovskii (voir [2, Théorème 4.2.1]) permet de prouver la première conséquence. L'application de ce théorème nécessite que la fonction objectif primale soit bornée inférieurement sur l'ensemble admissible primal. La satisfaction de cette exigence est déduite de [2, Lemme 4.2.1]. Par ailleurs, la seconde implication est démontrée dans [9].

Nous supposons donc, pour l'algorithme primal-dual, disposer de x_0 et de (y_0, s_0) des points strictement admissibles respectivement du primal et du dual qui constitueront l'itéré de départ de cet algorithme. Les algorithmes primaux utiliseront, quant à eux, un point x_0 strictement admissible du primal comme itéré de départ ainsi que ξ_0 une borne inférieure de la valeur optimale commune à (P) et à (D) notée ξ^*. Remarquons que l'existence de ξ_0 est garantie par [2,Lemme 4.2.1].

5.2 Etude des projections obliques

5.2.1 Définition des projections obliques

Définition 5.1 *Etant donnés ω un point intérieur de K fixé et u un élément de E^*, nous considérons le système linéaire en $(y(u),p(u))$ suivant:*

$$(5.1) \qquad \begin{cases} Ap(u) = 0 \\ A^*y(u) + F''(\omega)p(u) = u \end{cases}$$

La solution partielle $p(u)$ est appelée projection de u sur le noyau de A relativement à l'opérateur $F''(\omega)$.

Nous ferons également référence à $p(u)$ par la désignation plus succinte de projection oblique. Cette appellation s'appliquera aussi à $y(u)$.

5.2.2 Propriétés des projections obliques

Existence et unicité des projections obliques

Proposition 5.1 *Si A est un opérateur linéaire surjectif de E dans Y^*, alors le système (5.1) admet une solution unique.*

Preuve:
Nous montrons que l'opérateur linéaire défini par le premier membre de (5.1) est bijectif. Cet opérateur associe à un couple $(y(u), p(u))$ du produit cartésien $Y \times E$ le couple $(0, u)$ appartenant à $Y^* \times E^*$. Nous remarquons donc que son ensemble d'arrivée est le dual de son ensemble de définition. Ce dernier est de dimension finie. Par conséquent, l'opérateur considéré définit une application linéaire entre des espaces de dimensions finies. Démontrer son injectivité suffit dès lors à prouver son caractère bijectif. Par ailleurs, pour démontrer son injectivité, nous pouvons nous contenter, vu sa linéarité, de montrer que son noyau se réduit à l'élément nul. Lorsque u vaut zéro, la seconde équation du système (5.1) devient la suivante:

$$A^*y(0) + F''(\omega)p(0) = 0.$$

Nous considérons alors le produit scalaire de celle-ci avec $p(0)$:

$$< A^*y(0), p(0) > + < F''(\omega)p(0), p(0) > = < 0, p(0) > .$$

La première équation du système (5.1) évaluée en $u = 0$ permet de réduire l'égalité ci-dessus comme suit :

$$< F''(\omega)p(0), p(0) >= 0.$$

Le caractère défini positif imposé aux hessiens de F conclut alors à la nullité de $p(0)$. Dès lors, la seconde équation du système (5.1) en $u = 0$ devient la suivante :

$$A^*y(0) = 0.$$

Cette égalité suffit à entraîner l'annulation de $y(0)$. En effet, l'adjointe de A est linéaire et injective, étant données les hypothèses satisfaites par A.

□

Linéarité des projections obliques

Proposition 5.2 *La projection de u sur le noyau de A relativement à l'opérateur $F''(\omega)$ dépend linéairement de u.*

Ce résultat découle du principe de superposition associé aux systèmes linéaires ainsi que de l'unicité des projections obliques établie dans la Proposition 5.1.

Normes des projections obliques

Proposition 5.3 *Si $p(u)$ est la projection de u sur le noyau de A relativement à $F''(\omega)$, alors, nous avons l'égalité suivante :*

$$\|p(u)\|_\omega^2 =< u, p(u) > .$$

Cette assertion se déduit des définitions de $p(u)$ et de la norme $\|.\|_\omega$.

Proposition 5.4 *Etant donné $p(u)$ la projection de u sur le noyau de A relativement à $F''(\omega)$, nous avons, quel que soit y un élément de Y, l'inégalité suivante :*

$$\|p(u)\|_\omega \leq \|u - A^*y\|_\omega^*.$$

Cette inégalité devient une égalité seulement si y est le minimum du second membre.

Preuve:
Nous constatons que l'égalité est atteinte pour $y = y(u)$. Cette constatation découle de la première équation du système (5.1) ainsi que de l'égalité existant entre $\|F''(\omega)p(u)\|_\omega^*$ et $\|p(u)\|_\omega$.
Nous prouvons ensuite que $y(u)$ est le minimum unique de $\|u - A^*y\|_\omega^*$. Pour ce faire, nous démontrons, tout d'abord, que $y(u)$ est un extremum de cette fonction. Nous calculons son gradient :

$$\begin{aligned} \frac{d}{dy}\|u - A^*y\|_\omega^* &= \frac{d}{dy} < u - A^*y, [F''(\omega)]^{-1}(u - A^*y) > \\ &= -2A^{*^*}[F''(\omega)]^{-1}(u - A^*y). \end{aligned}$$

Nous remarquons que celui-ci s'annule en les y satisfaisant l'équation suivante :

$$A[F''(\omega)]^{-1}(u - A^*y) = 0.$$

$y(u)$ constitue un tel y du fait qu'il solutionne les équations décrites en (5.1). Nous établissons alors la stricte convexité de la fonction $\|u - A^*y\|_\omega^*$ considérée. Dans ce but, nous calculons son hessien :

$$\begin{aligned} \frac{d^2}{dy^2}\|u - A^*y\|_\omega^* &= \frac{d}{dy}(-2A[F''(\omega)]^{-1}(u - A^*y)) \\ &= 2A[F''(\omega)]^{-1}A^*. \end{aligned}$$

Nous constatons que celui-ci est défini positif. Cette constatation résulte du caractère défini positif imposé aux hessiens de F et suffit à garantir la stricte convexité de $\|u - A^*y\|_\omega^*$ qui admet, dès lors, un minimum global unique, $y(u)$.

□

Détermination des projections obliques

Dans cette section, nous détaillons et apprécions au point de vue calculatoire une méthode de détermination de la projection oblique $p(u)$, solution de (5.1). Ce système peut être réécrit comme suit :

$$(5.2) \quad \begin{cases} Ap(u) = 0 \\ -A^*y(u) + s(u) = 0 \\ s(u) + F''(\omega)p(u) = u \end{cases}$$

Nous obtenons, en combinant la deuxième et la troisième équation de ce nouveau système, l'égalité suivante :

$$A^*y(u) = u - F''(\omega)p(u).$$

En la pré-multipliant par $A[F''(\omega)]^{-1}$ et en tirant parti de la première équation du système (5.2), nous aboutissons au système linéaire en $y(u)$ suivant :

$$A[F''(\omega)]^{-1}A^*y(u) = A[F''(\omega)]^{-1}u$$

dont la résolution permet la détermination de $y(u)$. $s(u)$ et $p(u)$ peuvent alors, aux vues de la deuxième et de la troisième équation du système (5.2), être calculés successivement grâce aux formules suivantes :

$$\begin{aligned} s(u) &= A^*y(u) \\ p(u) &= [F''(\omega)]^{-1}(u - s(u)). \end{aligned}$$

Par conséquent la méthode de calcul de $p(u)$ établie est la suivante :

1. Calculer v défini par $v \stackrel{def}{=} [F''(\omega)]^{-1}u$.
2. Résoudre le système linéaire en $y(u)$ suivant :
$$A[F''(\omega)]^{-1}A^*y(u) = Av.$$
3. Calculer successivement $s(u)$ et $p(u)$ grâce aux formules ci-dessous :
$$\begin{aligned} s(u) &= A^*y(u) \\ p(u) &= [F''(\omega)]^{-1}(u - s(u)). \end{aligned}$$

La détermination de la projection oblique $p(u)$ nécessite donc l'inversion de $F''(\omega)$, la résolution d'un système linéaire et le calcul de produits de matrices par des vecteurs. Comme le hessien d'une barrière "ν-self-scaled" revêt généralement une forme relativement simple, l'inversion de $F''(\omega)$ ne pose pas de problème. D'autre part, la matrice du système linéaire à résoudre présente des propriétés dont la prise en compte, facilite grandement la résolution du système en question. Il s'agit d'une matrice carrée d'ordre m où m désigne la dimension de Y, qui, de plus, est symétrique et définie positive. Cette dernière caractéristique est induite par le caractère défini positif imposé aux hessiens de F. Par conséquent, le coût calculatoire de la détermination de la projection oblique $p(u)$ est raisonnable.

Comme annoncé dans l'introduction, nous verrons que les directions de recherche générées par les algorithmes étudiés dans le cadre de ce mémoire s'apparentent à des projections obliques. Par conséquent, toutes les propriétés établies ci-dessus pourront s'appliquer ces directions de recherche. Dès lors, nous disposons d'ores et déjà d'une méthode de calcul de celles-ci. Lors de l'établissement des méthodes de réduction de potentiel, nous nous contenterons donc d'expliciter l'obtention du système linéaire définissant les directions de recherche et nous n'aborderons pas la question de leurs déterminations.

Les chapitres restant sont consacrés à l'élaboration de deux méthodes de réduction de potentiel primales et d'une méthode de réduction de potentiel primale-duale symétrique.

Chapitre 6

Méthodes de réduction de potentiel primales

Dans ce chapitre, nous étendons les méthodes de réduction de potentiel primales établies par N. K. Karmarkar (voir [3], [8]) et C. C. Gonzaga (voir [4], [8]) pour la programmation linéaire au problème primal (P) décrit en 5.1. Yu. Nesterov et A. Nemirovskii ont déjà généralisé la méthode de Karmarkar à la programmation convexe conique (voir [2, chapitre 4]). L'algorithme qu'ils ont obtenu est à petits pas. Autrement dit, le pas réalisé à chaque itération progresse d'une fraction de la distance à la frontière de K mais reste confiné à l'intérieur de la boule unité centrée en l'itéré considéré et définie par la norme locale. Puisque la considération des barrières "ν-self-scaled" et des nouvelles distances introduites au Chapitre 3 a permis de lever un tel confinement dans l'établissement des bornes sur les hessiens et sur la variation de la barrière, une généralisation à grands pas des méthodes de Karmarkar et de Gonzaga paraît envisageable et devient l'objectif que nous nous efforcerons d'atteindre.

Les méthodes de réduction de potentiel primales que nous établirons utilisent la fonction potentiel primale paramétrisée par ξ, une borne inférieure de la valeur optimale primale, et définie sur l'ensemble primal strictement admissible par

$$\phi(x;\xi) \stackrel{def}{=} \mu \ln(< c, x > -\xi) + F(x) \tag{6.1}$$

où μ peut prendre différentes valeurs. Nous supposons satisfaites les hypothèses décrites en 5.1. Remarquons que celles-ci garantissent la disponibilité d'une borne inférieure ξ de la valeur optimale. De plus, nous admettons que la fonction objectif primale n'est jamais constante sur la région admissible primale. Ainsi, la valeur de cette fonction en tout point strictement admissible du primal excède la valeur optimale. Cette hypothèse supplémentaire assure donc que la fonction potentiel primale $\phi(x;\xi)$ est bien définie sur l'ensemble primal strictement admissible. C'est là son unique utilité. Signalons d'ailleurs son caractère non restrictif. En effet, si elle n'est pas satisfaite, il suffit d'adjoindre aux algorithmes que nous déterminerons un critère d'arrêt supplémentaire qui, lorsque ces algorithmes disposent d'un point admissible x et d'une borne inférieure ξ tels que $< c, x >= \xi$, termine ceux-ci en précisant l'optimalité de x.

Ce chapitre est divisé en deux parties. La première est consacrée à l'extension de la méthode de Karmarkar et la seconde à la généralisation de l'algorithme de Gonzaga. Ces deux parties partagent la même structure, exception faite du premier point traité dans l'extension de la méthode de Karmarkar. Nous dégageons, tout d'abord, les conditions que doivent satisfaire nos extensions pour converger vers l'optimalité primale. Nous détaillons, ensuite, l'obtention de celles-ci en élaborant des procédures et en établissant des résultats qui garantissent la satisfaction de ces conditions de convergence. Finalement, nous apprécions les algorithmes ainsi obtenus.

6.1 Extension de la méthode de Karmarkar

6.1.1 Formes particulières du problème traité et de la fonction potentiel primale utilisée

La méthode de Karmarkar établie pour la programmation linéaire ne traite pas le problème primal sous sa forme standard (voir [8], page 15). Similairement, notre extension de celle-ci ne résoudra pas directement le problème primal (P) décrit en 5.1. En substituant $E \times I\!R$ à E, (x, τ) à x, $K \times I\!R^+$ à K et $Ax - b\tau = 0$ où $\tau = 1$ à $Ax = b$, nous réécrivons, sans perte de généralité, la contrainte $Ax = b$ du problème primal (P) décrit en 5.1 comme suit :

$$\begin{array}{l} Bx = 0 \\ < d, x >= 1 \end{array}$$

où d appartient à K^*. Si nous désignons par $\tilde{Y}^*$ l'image de B, nous obtenons que $Y^* = \tilde{Y}^* \times I\!R$. Par ailleurs, nous désignons par le couple (y, ξ) appartenant à $\tilde{Y} \times I\!R$ un élément de Y. Les problèmes primal et dual que traitera notre extension de l'algorithme de Karmarkar, sont respectivement :

$$\begin{array}{l} \min < c, x > \\ sc. \\ \quad Bx = 0 \\ \quad < d, x >= 1 \\ \quad x \in K \end{array}$$

où d appartient à K^* et :

$$\begin{array}{l} \max \xi \\ sc. \\ \quad B^*y + d\xi + s = c \\ \quad s \in K^* \end{array}$$

où B^* désigne l'adjointe de B. y est dès lors un élément de $\tilde{Y}$. Dans la suite de cette section, nous désignerons respectivement par (P) et (D) les problèmes primal et dual mis sous la forme décrite ci-dessus.

Puisque les hypothèses répertoriées en 5.1 sont supposées satisfaites, le problème dual (D) admet une solution optimale notée (y^*, ξ^*, s^*). La valeur optimale commune au primal et au dual vaut dès lors ξ^*.

D'autre part, notre généralisation de la méthode de Karmarkar utilise la fonction potentiel primale décrite en (6.1) dans laquelle μ se voit attribué la valeur ν. A l'origine, le domaine de définition de cette fonction se limite à l'ensemble primal strictement admissible. Nous étendons celui-ci à $\{x \in \text{ int } K \;:\; Bx = 0\}$. Pour ce faire, nous définissons la fonction potentiel primale comme suit :

$$\phi(x;\xi) \stackrel{def}{=} \nu \ln < c - \xi d, x > + F(x).$$

Remarquons que celle-ci coïncide bien avec celle définie en (6.1) sur l'ensemble primal strictement admissible. Par ailleurs, l'attribution de la valeur ν à μ entraîne l'homogénéité de degré zéro de la fonction potentiel primale considérée. Cela signifie que l'égalité ci-dessous est satisfaite pour tout t strictement positif :

$$\phi(tx;\xi) = \phi(x;\xi).$$

Cette propriété découle de l'homogénéité ν-logarithmique de F, caractéristique d'une barrière ν-normale. Cet attribut de la fonction potentiel primale facilitera la détermination de l'itéré suivant dans notre généralisation de la méthode de Karmarkar. Cette facilitation que nous détaillerons dans le point 6.1.3, avait déjà motivé l'extension de la définition de la fonction potentiel primale à $\{x \in \text{ int } K \;:\; Bx = 0\}$.

6.1.2 Détermination des conditions de convergence vers l'optimalité primale

Proposition 6.1 *La barrière "ν-self-scaled" F considérée est bornée inférieurement sur l'ensemble $\bar{K}$ défini par*

$$\bar{K} \stackrel{def}{=} \{x \in K : \; Bx = 0, \; < c, x > \leq \gamma_0, \; < d, x > \; \leq \; 1\}$$

où $\gamma_0 \stackrel{def}{=} \max(< c, x_0 >, 0) + 1$.

Preuve:

1. Nous montrons que l'ensemble $\bar{K}$ est borné. Pour ce faire, nous démontrons, tout d'abord, le caractère borné de l'ensemble $\bar{K}_1$ défini par

$$\bar{K}_1 \stackrel{def}{=} \{x \in K : \ Bx = 0, \ < c, x > \leq \gamma_0, \ < d, x >= 1\}.$$

Tout problème de minimisation sous contraintes peut, grâce à l'introduction de la fonction de support de son ensemble admissible, se ramener à un problème de minimisation sans contraintes. Ainsi, le problème primal (P) considéré est équivalent au problème de minimisation convexe sans contraintes suivant :

$$\min f(x)$$

où $f(x) \stackrel{def}{=} < c, x > +\psi_C(x)$, ψ_C désignant la fonction d'appui de l'ensemble admissible primal noté C. Puisque nous supposons satisfaites les hypothèses décrites en 5.1, ξ^* désigne la valeur optimale commune à (P) et à (D) et nous obtenons également que l'ensemble des solutions primales

$$\{x \ : \ f(x) \leq \xi^*\}$$

est non vide et borné. Désireux de prouver le caractère borné de $\bar{K}_1$, nous utilisons un résultat d'analyse convexe établi par R.T. Rockafellar (voir [6, Corollaire 8.7.1]) et déja énoncé dans la preuve du Théorème 2.2 sous la forme du Corollaire 1. Notre volonté d'application de celui-ci nous contraint à vérifier que f est convexe, propre et fermée. La convexité étant acquise, nous nous attachons à démontrer le caractère propre de f. On vérifie aisément que $f(x)$ excède strictement $-\infty$ en tout x point du domaine de définition de f. D'autre part, le point strictement admissible du primal x_0 est tel que $f(x_0)$ est strictement inférieur à $+\infty$. La fermeture de f découle, quant à elle, de la continuité du produit scalaire et de la fermeture de C. Par conséquent, nous obtenons, par application du Corollaire 1, le caractère borné de l'ensemble suivant :

$$\{x \ : \ f(x) \leq \alpha\}$$

et ce, quel que soit le réel α considéré. La particularisation de ce résulat à $\alpha = \gamma_0$ conclut au caractère borné de $\bar{K}_1$. Nous établissons, ensuite, ce même caractère pour l'ensemble $\bar{K}_0$ défini par

$$\bar{K}_0 \stackrel{def}{=} \{x \in K : \ Bx = 0, \ < c, x > \leq \gamma_0, \ < d, x >= 0\}.$$

A nouveau, nous avons recours à un résultat d'analyse convexe établi par R.T. Rockafellar (voir [6,Corollaire 8.4.1]).

Corollaire 2 *Soient C une partie convexe fermée et M un ensemble affine dont l'intersection avec C est non vide et bornée.*
Alors, l'intersection de C avec tout autre ensemble affine M' parallèle à M, est bornée.

Nous appliquons le Corollaire 2 avec les identifications $C = K$ et $M = \{x : Bx = 0, < c, x > \leq \gamma_0, < d, x >= 1\}$. Le résultat démontré en premier lieu et le fait que γ_0 soit, par définition, strictement plus grand que ξ^* permettent la vérification des hypothèses imposées à l'intersection de C et de M. Nous obtenons dès lors l'assertion escomptée par application du Corollaire 2 à $M' = \{x : Bx = 0, < c, x > \leq \gamma_0, < d, x >= 0\}$. Nous déduisons, finalement, le caractère borné de $\bar{K}$. Nous désignons respectivement par M_0 et M_1 les bornes de $\bar{K}_0$ et de $\bar{K}_1$, et nous montrons, pour tout x appartenant à $\bar{K}$, l'inégalité suivante :

$$\|x\| \leq M_0 + M_1.$$

Remarquons que celle-ci est immédiatement satisfaite par les éléments de $\bar{K}$ dont le produit scalaire avec d vaut 0 ou 1. Nous considérons, dès lors, x un élément arbitraire de $\bar{K}$, tel que

$$< d, x >= t$$

où t est un réel strictement compris entre 0 et 1. Etant donné y un élément de $\bar{K}_1$, nous définissons z par

$$z \stackrel{def}{=} \frac{x - ty}{1 - t}.$$

On vérifie facilement l'appartenance de z à $\bar{K}_0$. Dès lors, le développement suivant :

$$\begin{aligned} \|x\| &= \|ty + (1-t)z\| \\ &\leq t\|y\| + (1-t)\|z\| \\ &\leq \|y\| + \|z\| \\ &\leq M_1 + M_0 \end{aligned}$$

permet bien de conclure au caractère borné de $\bar{K}$.

2. Nous montrons que F est bornée inférieurement sur $\bar{K}$. La fermeture de K ainsi que la continuité du produit scalaire et de l'opérateur linéaire B entraînent la fermeture de $\bar{K}$. D'autre part, F étant notamment une barrière ν-logarithmiquement homogène pour K, elle est de classe C^2 et donc, continue. Par conséquent, ces deux constatations combinées au caractère borné de $\bar{K}$ établi en premier lieu, impliquent le résultat escompté, par le théorème de Weierstrass.

□

Théorème 6.1 *Soient $\underline{F}$ le minimum de F sur $\bar{K}$, x un point strictement admissible du primal et ξ une borne inférieure de la valeur optimale supérieure ou égale à ξ_0 et telle que*

$$\phi(x; \xi) \leq \phi(x_0; \xi_0) - \Delta.$$

Alors, tant que la diminution de la fonction potentiel primale Δ *satisfait la condition suivante :*

$$\Delta > \nu \ln(< c, x_0 > -\xi_0) + F(x_0) - \underline{F}, \tag{6.2}$$

nous avons l'inégalité ci-dessous :

$$< c, x > -\xi \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0) - \underline{F}}{\nu}} e^{\frac{-\Delta}{\nu}} \tag{6.3}$$

Preuve:

Remarquons que l'existence de $\underline{\mathrm{F}}$ est garantie par la Proposition 6.1. Nous définissons d'une part,

$$\lambda \stackrel{def}{=} \max(1, \frac{< c, x >}{\gamma_0})$$

où $\gamma_0 \stackrel{def}{=} \max(< c, x_0 >, 0) + 1$ et d'autre part,

$$\bar{x} \stackrel{def}{=} \frac{x}{\lambda}.$$

Nous distinguons alors deux cas. Soit le point strictement admissible considéré x est un élément de $\bar{K}$. Cette appartenance implique que

$$\lambda = 1$$

et donc que $\bar{x}$ coïncide avec x. Soit le point strictement admissible x n'appartient pas à $\bar{K}$. Dans ce cas, λ se voit attribué la valeur suivante :

$$\frac{< c, x >}{\gamma_0}$$

et $\bar{x}$ qui vaut $\frac{x}{\lambda}$, est dès lors un élément de $\bar{K}$ tel que $< c, \bar{x} >= \gamma_0$. L'homogénéité de degré zéro de la fonction ϕ implique, dans les deux cas, l'égalité ci-dessous :

$$\phi(\bar{x}, \xi) = \phi(x, \xi).$$

Dès lors, puisque nous supposons que $\phi(x; \xi) \leq \phi(x_0; \xi_0) - \Delta$, nous obtenons, par explicitation de ϕ, l'inégalité ci-dessous :

$$\nu \ln < c - \xi d, \bar{x} > + F(\bar{x}) \leq \nu \ln < c - \xi d, x_0 > + F(x_0) - \Delta.$$

En divisant celle-ci par le réel strictement positif ν et en appliquant à ces deux membres la fonction exponentielle, nous aboutissons à la relation suivante :

$$< c - \xi d, \bar{x} > \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0) - F(\bar{x})}{\nu}} e^{\frac{-\Delta}{\nu}}.$$

Puisque, dans les deux cas, le point $\bar{x}$ appartient à $\bar{K}$, l'inégalité ci-dessous est satisfaite :

$$F(\bar{x}) \geq \underline{F}$$

et par suite, la suivante également :

$$e^{\frac{-F(\bar{x})}{\nu}} \leq e^{\frac{-\underline{F}}{\nu}}.$$

Par conséquent, nous pouvons majorer le second membre de l'inégalité obtenue précédemment par une expression indépendante de $\bar{x}$ et obtenir de la sorte la relation suivante :

$$< c - \xi d, \bar{x} > \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0)-\underline{F}}{\nu}} e^{\frac{-\Delta}{\nu}}. \tag{6.4}$$

Nous nous intéressons, tout d'abord, au second cas et démontrons le non respect de la condition (6.2) sous cette alternative. L'appartenance de x à l'ensemble primal strictement admissible, la définition de γ_0, le fait que la valeur de la fonction objectif primale en tout point strictement admissible excède la valeur optimale et les contraintes satisfaites par ξ légitiment le développement suivant :

$$\begin{aligned} < c - \xi d, \bar{x} > &= \frac{1}{\lambda} < c, x > - \frac{\xi}{\lambda} < d, x > \\ &= \frac{1}{\lambda} < c, x > - \frac{\xi}{\lambda} \\ &= \gamma_0 - \frac{\xi}{\lambda} \\ &\geq 1 + \frac{< c, x_0 >}{\lambda} - \frac{\xi}{\lambda} \\ &\geq 1. \end{aligned}$$

La minoration du premier membre de (6.4) ainsi obtenue nous conduit à l'inégalité suivante :

$$1 \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0)-\underline{F}}{\nu}} e^{\frac{-\Delta}{\nu}}.$$

En appliquant le logarithme népérien aux deux membres de celle-ci et en multipliant ceux-ci par le réel strictement positif ν, nous aboutissons à la relation ci-dessous :

$$0 \leq \nu \ln(< c, x_0 > -\xi_0) + F(x_0) - \underline{F} - \Delta$$

qui est exactement la négation de la condition (6.2) imposée à la diminution de la fonction potentiel primale Δ. Ce résultat intermédiaire établit donc que la seconde alternative ne vérifie pas toutes les hypothèses du Théorème 6.1. Par conséquent, nous pouvons nous contenter de démontrer celui-ci uniquement pour un point strictement admissible du primal appartenant à $\bar{K}$. Comme nous l'avons vu, cela entraîne que x égale $\bar{x}$. L'inégalité (6.4) devient dès lors la suivante :

$$< c, x > -\xi < d, x > \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0)-\underline{F}}{\nu}} e^{\frac{-\Delta}{\nu}}.$$

L'appartenance de x à l'ensemble primal admissible qui implique l'égalité suivante :

$$< d, x >= 1$$

permet alors d'aboutir à l'inégalité (6.3).

□

Corollaire 6.1 *Supposons que nous disposions d'un point strictement admissible du primal x_0, d'une borne inférieure de la valeur optimale ξ_0 et d'un algorithme qui, à chaque itération, remplace l'itéré strictement admissible du primal x_k et la borne inférieure de la valeur optimale ξ_k respectivement par x_{k+1} et ξ_{k+1} de même nature et ce, de telle façon que*

$$\phi(x_{k+1}; \xi_{k+1}) \leq \phi(x_k; \xi_k) - \delta$$

où δ est strictement positif.
Alors, pour un ϵ suffisamment petit (en fait, strictement inférieur à 1), nous atteignons l'ϵ-optimalité primale après K itérations, où

$$K = \lceil \frac{\nu}{\delta}(\ln(< c, x_0 > -\xi_0) + \frac{F(x_0) - \underline{F}}{\nu} + \ln(\frac{1}{\epsilon})) \rceil.$$

<u>**Preuve:**</u>
Remarquons que la démonstration de l'assertion suivante :

$$< c, x_K > -\xi_K \leq \epsilon$$

suffit à garantir l'accession à l'ϵ-optimalité primale après K itérations. Cette suffisance s'explique par la nature de ξ_K. Nous constatons que l'hypothèse relative à la diminution de la fonction potentiel primale implique l'inégalité suivante :

$$\phi(x_K; \xi_K) \leq \phi(x_0; \xi_0) - K\delta.$$

Nous appliquons alors le Théorème 6.1 avec les identifications $x = x_K$, $\xi = \xi_K$ et $\Delta = K\delta$. Le fait que ϵ soit supposé strictement inférieur à 1 entraîne la satisfaction de la condition (6.2) par la diminution de la fonction potentiel primale $K\delta$. Nous obtenons dès lors, par application du Théorème 6.1, l'inégalité suivante :

$$< c, x_K > -\xi_K \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0) - \underline{F}}{\nu}} e^{\frac{-K\delta}{\nu}}.$$

La définition de K nous permet de majorer le second membre de celle-ci comme suit :

$$(< c, x_0 > -\xi_0) e^{\frac{F(x_0) - \underline{F}}{\nu}} e^{\frac{-\delta}{\nu}[\frac{\nu}{\delta}(\ln(< c, x_0 > -\xi_0) + \frac{F(x_0) - \underline{F}}{\nu} + \ln(\frac{1}{\epsilon}))]}$$

expression qui, après simplification, se réduit à bien ϵ.

□

Le Théorème 6.1 montre qu'une diminution suffisante de la fonction potentiel primale considérée entraîne un rapprochement de l'optimalité primale. Notre extension de la méthode de Karmarkar doit donc s'efforcer de diminuer la valeur de sa fonction potentiel au cours de ses itérations.

Le Corollaire 6.1 exploite ce résultat afin d'établir les conditions à satisfaire pour atteindre l'ϵ-optimalité primale ainsi que le nombre d'itérations nécessaires à l'obtention de celle-ci. Il nous apprend, en effet, qu'il suffit que la fonction potentiel primale décroisse à chaque itération d'une constante suffisante pour obtenir, après $O(\nu \ln(\frac{1}{\epsilon}))$ itérations, une solution primale ϵ-optimale . Soulignons aussi la souplesse des méthodes de réduction de potentiel primales qui autorisent une mise à jour de la borne inférieure utilisée. Cette possibilité facilite évidemment la garantie d'une diminution constante et suffisante de la fonction potentiel primale à chaque itération. Par ailleurs, nous sommes conscients d'avoir négligé des quantités dans l'établissement de notre borne de complexité. Cette négligence s'explique, d'une part, par un souci de généralité. En effet, les quantités dont nous n'avons pas tenu compte, dépendent du problème (P) considéré et de l'initialisation. D'autre part, nous supposons que ϵ est pris suffisamment petit que pour rendre ces quantités négligeables par rapport à $\frac{\nu}{\rho} \ln(\frac{1}{\epsilon})$.

6.1.3 Mise en oeuvre de procédures assurant la satisfaction des conditions de convergence

Supposons, comme dans le Corolloraire 6.1, qu'à chaque itération de notre extension de la méthode de Karmarkar, nous disposions d'un point strictement admissible du primal $\hat{x}$ et d'une borne inférieure de la valeur optimale $\hat{\xi}$. La question à laquelle nous répondons dans ce point est celle de la détermination de l'itéré suivant et de la mise à jour éventuelle de la borne inférieure de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 6.1 et le Corollaire 6.1.

Soulignons, avant toute chose, de quelle façon l'homogénéité de degré zéro de la fonction potentiel primale facilite la détermination du prochain itéré. Celui-ci est soumis aux deux contraintes suivantes. Il doit, d'une part, fournir une réduction suffisante de la fonction potentiel primale afin de garantir la convergence de la méthode et d'autre part, appartenir à l'ensemble primal strictement admissible pour permettre les itérations de l'algorithme. Nous pouvons cependant nous contenter de déterminer un itéré $\tilde{x}_+$ qui vérifie la première contrainte, satisfait à $B\tilde{x}_+ = 0$ et appartient à l'intérieur de K. Autrement dit, nous ne tenons pas compte de la normalisation $< d, \tilde{x}_+ >= 1$ incluse dans la seconde contrainte. Notons que cette négligence nécessite et donc justifie l'extension de la définition de la fonction potentiel primale à $\{x \in \text{ int } K \ : \ Bx = 0\}$ réalisée en 6.1.1.

L'itéré effectif x_+ est alors obtenu comme suit :

$$x_+ = \frac{\tilde{x}_+}{< d, \tilde{x}_+ >}.$$

On montre aisément sa stricte admissibilité primale. Par ailleurs, il engendre une diminution de la fonction potentiel primale semblable à celle fournie par $\tilde{x}_+$. Cette conservation de la diminution obtenue s'explique par l'homogénéité de degré zéro de la fonction potentiel primale. Par conséquent, l'intérêt de cette propriété est de nous permettre, lors de la détermination de l'itéré suivant, de négliger la contrainte de normalisation.

Détermination de la direction de descente

Puisque nous voulons, à chaque itération, diminuer la valeur de la fonction potentiel primale, nous devons calculer une direction de descente de ϕ en $\hat{x}$. Nous recherchons cette direction dans le noyau de B. La présence de cette contrainte s'explique par notre volonté de générer un prochain itéré qui soit strictement admissible pour le primal. Il n'est pas nécessaire d'introduire les deux autres contraintes primales. En effet, le respect de la contrainte d'appartenance à l'intérieur de K est pris en charge par la fonction barrière F et nous avons vu ci-dessus que nous pouvions négliger la contrainte de normalisation. Par ailleurs, nous préférons à la méthode de Newton une méthode de quasi-Newton. Celle-ci nous permet d'approximer en $\hat{x}$ le hessien de ϕ par celui de F qui présente l'avantage d'être défini positif. Dès lors, le problème de minimisation sous contraintes obtenu :

$$\begin{array}{ll} \min\limits_{p} & < u, p > + \frac{1}{2} < F''(\hat{x})p, p > \\ sc. & Bp = 0 \end{array}$$

où $u \stackrel{def}{=} \phi'(\hat{x}; \xi)$ et où ξ désigne une borne inférieure de la valeur optimale éventuellement mise à jour, est convexe. Nous désignons par $-p(u)$ la solution de ce problème. La convexité tant de son ensemble admissible que de sa fonction objectif nous permet de le réécrire, de façon équivalente, sous la forme d'un système de Karush-Kühn-Tücker. De la sorte, la direction de recherche $p(u)$ est obtenue en résolvant le système linéaire en $(y(u), p(u))$ suivant :

$$\begin{array}{rcl} B^*y(u) + F''(\hat{x})p(u) & = & u \\ Bp(u) & = & 0 \end{array} \tag{6.5}$$

où $y(u)$ désigne le vecteur des multiplicateurs de Lagrange associé à la contrainte d'égalité $Bp = 0$. L'existence et l'unicité de la direction de recherche $p(u)$ sont garanties par la Proposition 5.1 puisque, comme annoncé, nous constatons que la direction de recherche générée s'apparente à une projection oblique. De plus, cette constatation nous autoriserait à ne pas aborder la question de la résolution du système linéaire (6.5). Cependant, celui-ci n'est pas toujours immédiatement résolvable. Cette éventuelle impossibilité de résolution du système (6.5) provient de l'intervention de ξ dans celui-ci. En effet, u dépend de ξ. Cette borne inférieure de la valeur optimale peut avoir été mise à jour. Dans ce cas, sa valeur nous est momentanément inconnue et nous ne savons donc pas —du moins, pas complètement— résoudre le système (6.5). Dès lors, nous optons pour une méthode de résolution en deux temps dont le premier est indépendant de la valeur de ξ. Dans ce but, nous détaillons l'obtention d'une équation liant $p(u)$ à ξ et dont ces quantités sont les deux seules inconnues. Nous explicitons u et obtenons de la sorte l'égalité suivante :

$$u = \frac{\nu}{< c - \xi d, \hat{x} >}(c - \xi d) + F'(\hat{x}).$$

Nous définissons ensuite les deux quantités ci-dessous :

$$\begin{aligned} \hat{c} &\stackrel{def}{=} c + \frac{< c, \hat{x} >}{\nu} F'(\hat{x}) \\ \hat{d} &\stackrel{def}{=} d + \frac{< d, \hat{x} >}{\nu} F'(\hat{x}) \end{aligned}$$

qui sont indépendantes de ξ et qui nous permettent d'exprimer u comme suit :

$$u = \frac{\nu}{< c - \xi d, \hat{x} >}(\hat{c} - \xi \hat{d}).$$

Comme la Proposition 5.2 établit la dépendance linéaire de la direction de recherche $p(u)$ par rapport à u, nous aboutissons à l'équation suivante :

$$p(u) = \frac{\nu}{< c - \xi d, \hat{x} >}(p(\hat{c}) - \xi p(\hat{d}))$$

qui rencontre bien nos exigences. Remarquons la positivité du rapport $\frac{\nu}{<c-\xi d, \hat{x}>}$. La direction de recherche $p(u)$ est donc un multiple positif de $p(\hat{c}) - \xi p(\hat{d})$. Les projections obliques $p(\hat{c})$ et $p(\hat{d})$ sont définies par des systèmes linéaires indépendants de ξ et peuvent dès lors être immédiatement déterminées grâce à la méthode décrite à la fin du Chapitre 5. Ces déterminations font donc l'objet de la première phase de notre méthode de résolution du système (6.5). La seconde, quant à elle, aura lieu une fois la valeur de ξ connue, c'est-à-dire après la mise à jour de la borne inférieure. Elle résoudra l'équation liant $p(u)$ à ξ, $p(u)$ constituant alors l'unique inconnue de cette équation.

Mise à jour de la borne inférieure

La première partie de cette subdivision est consacrée à la description de la mise à jour de la borne inférieure. Le résultat que nous établissons ensuite, légitime l'emploi de celle-ci lors des itérations ultérieures et souligne la pertinence de la considération des nouvelles distances introduites au Chapitre 3 dans les méthodes de réduction de potentiel primales. Finalement, nous détaillons l'obtention d'égalités qui seront utiles par la suite.

Afin de déterminer la mise à jour de $\hat{\xi}$, nous définissons, pour tout ξ borne inférieure de la valeur optimale, les deux quantités suivantes :

$$\begin{aligned} \tilde{x}(\xi) &\stackrel{def}{=} \frac{< c - \xi d, \hat{x} >}{\nu} \hat{x} + (p(\hat{c}) - \xi p(\hat{d})), \\ \tilde{s}(\xi) &\stackrel{def}{=} F''(\hat{x})\tilde{x}(\xi). \end{aligned}$$

Notons que l'explicitation de $\tilde{x}(\xi)$ et l'emploi de la première égalité décrite en (1.2) dans la définition de $\tilde{s}(\xi)$, permettent d'exprimer cette dernière quantité comme suit :

$$\tilde{s}(\xi) = \frac{- < c - \xi d, \hat{x} >}{\nu} F'(\hat{x}) + F''(\hat{x})(p(\hat{c}) - \xi p(\hat{d})). \tag{6.6}$$

Soulignons également un résultat sur lequel se basera la preuve du Lemme 6.1. En combinant les secondes équations des systèmes définissant $p(\hat{c})$ et $p(\hat{d})$, et en utilisant l'égalité (6.6) ainsi que les définitions de $\hat{c}$ et de $\hat{d}$, nous aboutissons à l'égalité suivante :

$$B^*(y(\hat{c}) - \xi y(\hat{d})) + \tilde{s}(\xi) = c - \xi d.$$

Par conséquent, tant que $\tilde{s}(\xi)$ appartient au cône dual K^*, le couple $(y(\hat{c}) - \xi y(\hat{d}), \tilde{s}(\xi))$ est un point admissible du dual (D). La fonction objectif duale évaluée en ce point vaut ξ qui constitue dès lors, lorsque l'appartenance de $\tilde{s}(\xi)$ à K^* est satisfaite, une borne inférieure de la valeur optimale commune à (P) et à (D).

L'éventualité de la mise à jour de la borne inférieure $\hat{\xi}$ dépend de l'appartenance ou non de $\tilde{x}(\hat{\xi})$ à l'intérieur de K. Lorsque $\tilde{x}(\hat{\xi})$ ne lui appartient pas, la mise à jour n'a pas lieu. Dès lors, nous attribuons à ξ_+, la borne inférieure associée à l'itération suivante, la valeur de la borne inférieure effective, à savoir $\hat{\xi}$. Par contre, dans le cas où $\tilde{x}(\hat{\xi})$ est un point intérieur de K, nous remarquons que la stricte admissibilité primale de $\hat{x}$ permet d'exprimer $\tilde{x}(\xi)$ comme suit :

$$\tilde{x}(\xi) = \tilde{x}(\hat{\xi}) - (\xi - \hat{\xi})\tilde{p}$$

où $\tilde{p} \stackrel{def}{=} p(\hat{d}) + \frac{\hat{x}}{\nu}$. Nous définissons alors ξ_+ de la façon suivante :

$$\xi_+ \stackrel{def}{=} \hat{\xi} + \frac{1}{\sigma_{\tilde{x}(\hat{\xi})}(\tilde{p})}.$$

Le but de la mise à jour est de fournir, lorsque cela est possible, une borne inférieure de la valeur optimale plus fine que celle utilisée à l'itération précédente. La mise à jour décrite ci-dessus atteint au moins partiellement cet objectif. En effet, la borne inférieure associée à l'itération suivante n'est jamais plus petite que la borne inférieure effective $\hat{\xi}$.

Par ailleurs, que $\tilde{x}(\hat{\xi})$ appartienne ou non à l'intérieur de K, nous pouvons exprimer la direction de recherche $p(u)$ comme un multiple positif de p, p étant défini par :

$$p \stackrel{def}{=} p(\hat{c}) - \xi_+ p(\hat{d}).$$

Lemme 6.1 *La mise à jour de la borne inférieure $\hat{\xi}$ notée ξ_+ satisfait les deux inégalités suivantes :*

(i.) $\xi_+ \leq \xi^*$

(ii.) $< c, \hat{x} > -\xi_+ \leq \nu \sigma_{\hat{x}}(-p)$ *où* $p \stackrel{def}{=} p(\hat{c}) - \xi_+ p(\hat{d})$.

Preuve:
Les preuves des deux assertions décrites dans le Lemme 6.1 requièrent la démonstration de l'appartenance de $\tilde{x}(\xi_+)$ à la frontière de K lorsque $\tilde{x}(\hat{\xi})$ est un point intérieur de K. Nous établissons donc ce résultat. Les définitions de ξ_+, $\tilde{x}(\xi)$ et de $\tilde{p}$ justifient le développement suivant :

$$\begin{aligned}
\tilde{x}(\xi_+) &= \tilde{x}(\hat{\xi} + \frac{1}{\sigma_{\tilde{x}(\hat{\xi})}(\tilde{p})}) \\
&= \frac{< c, \hat{x} > -\hat{\xi}}{\nu}\hat{x} + (p(\hat{c}) - \hat{\xi}p(\hat{d})) - \frac{1}{\sigma_{\tilde{x}(\hat{\xi})}(\tilde{p})}[\frac{\hat{x}}{\nu} + p(\hat{d})] \\
&= \tilde{x}(\hat{\xi}) - \frac{1}{\sigma_{\tilde{x}(\hat{\xi})}(\tilde{p})}\tilde{p}.
\end{aligned}$$

Le résultat escompté se déduit alors de la définition de $\sigma_{\tilde{x}(\hat{\xi})}(\tilde{p})$.

(*i.*) Cette assertion nécessite une démonstration uniquement si une mise à jour de la borne inférieure a eu lieu. Nous supposons donc que $\tilde{x}(\hat{\xi})$ est un point intérieur de K. Nous avons établi ci-dessus que l'appartenance de $\tilde{s}(\xi_+)$ au cône dual K^* implique que ξ_+ est une borne inférieure de la valeur optimale. Nous nous attachons dès lors à montrer cette appartenance. Celle-ci découle de la définition de $\tilde{s}(\xi_+)$, de l'appartenance de $\tilde{x}(\xi_+)$ à K démontrée ci-dessus et du point (*iii.*) du Théorème 2.1.

(*ii.*) Nous envisageons, tout d'abord, le cas où $\tilde{x}(\hat{\xi})$ n'est pas un point intérieur de K. Nous obtenons, grâce aux définitions de $\tilde{x}(\xi)$, de ξ_+ et de p ainsi qu'à la stricte admissibilité primale de $\hat{x}$, la suite d'égalités ci-dessous :

$$\begin{aligned}
\tilde{x}(\hat{\xi}) &= \frac{< c, \hat{x} > -\xi_+}{\nu}\hat{x} + (p(\hat{c}) - \xi_+ p(\hat{d})) \\
&= \frac{< c, \hat{x} > -\xi_+}{\nu}\hat{x} - (-p).
\end{aligned}$$

Par hypothèse, le dernier membre de ces égalités n'appartient pas à l'intérieur de K. Nous obtenons dès lors, par définition de $\sigma_{\hat{x}}(-p)$, l'inégalité suivante :

$$\frac{< c, \hat{x} > -\xi_+}{\nu} \leq \sigma_{\hat{x}}(-p).$$

La stricte positivité de ν conclut au résultat escompté.
Nous considérons ensuite le cas où $\tilde{x}(\hat{\xi})$ appartient à l'intérieur de K. Nous avons initialement montré l'appartenance de $\tilde{x}(\xi_+)$ à la frontière de K. Nous développons cette quantité de la même façon que nous avons explicité $\tilde{x}(\hat{\xi})$ ci-dessus. Nous concluons, à nouveau, par définition de $\sigma_{\hat{x}}(-p)$, à l'égalité suivante :

$$\frac{< c, \hat{x} > -\hat{\xi}}{\nu} = \sigma_{\hat{x}}(-p)$$

qui conduit bien au résultat attendu.

□

La première partie du Lemme 6.1 confirme que la mise à jour de la borne inférieure est une borne inférieure de la valeur optimale et légitime donc l'emploi de celle-ci lors des itérations ultérieures. Le point $(ii.)$ fournit, quant à lui, une borne qui permettra la démonstration d'un théorème qui garantit, à chaque itération, une diminution constante de la fonction potentiel primale. Nous remarquons que, si la norme locale définie par le hessien de la barrière était la seule mesure de la distance à la frontière de K disponible, la borne fournie aurait été plus grossière. Dès lors, la diminution constante de la fonction potentiel primale, assurée dans ce cas à chaque itération, aurait été moindre que celle garantie lorsque la nouvelle mesure $\sigma_x(.)$ est utilisée. Par conséquent, ce lemme justifie l'introduction des nouvelles distances à la frontière de K réalisée au Chapitre 3. Par ailleurs, nous déduisons du point $(ii.)$ l'implication suivante : si l'itéré $\hat{x}$ n'est pas optimal, alors $\sigma_{\hat{x}}(-p)$ est non nul.

La fin de cette subdivision est, comme annoncé, consacrée à l'obtention d'égalités dont nous nous servirons ultérieurement. La linéarité des projections obliques établie dans la Proposition 5.2, implique que p est la projection de $\hat{c} - \xi_+\hat{d}$ sur le noyau de B relativement à $F''(\hat{x})$. Dès lors, nous obtenons, par application de la Proposition 5.3 à p, la relation ci-dessous :

$$\|p\|_{\hat{x}}^2 =< \hat{c} - \xi_+\hat{d}, p > . \tag{6.7}$$

Nous montrons la nullité de $< F'(\hat{x}), p >$. Pour ce faire, nous considérons, tout d'abord, le produit scalaire de la première équation du système (6.5) avec $\hat{x}$. Nous obtenons de la sorte l'égalité suivante :

$$< B^*y(u), \hat{x} > + < F''(\hat{x})p(u), \hat{x} > = < u, \hat{x} > . \tag{6.8}$$

Nous développons ensuite le premier membre de celle-ci grâce à l'hypothèse de symétrie imposée aux hessiens de F, à la stricte admissibilité primale de $\hat{x}$ ainsi qu'à la première égalité décrite en (1.2). Nous obtenons ainsi la suite d'égalités ci-dessous :

$$\begin{aligned} < B^* y(u), \hat{x} > + < F''(\hat{x}) p(u), \hat{x} > &= < B\hat{x}, y(u) > + < F''(\hat{x})\hat{x}, p(u) > \\ &= - < F'(\hat{x}), p(u) > . \end{aligned}$$

Nous déduisons alors la nullité du second membre de (6.8) pour $u = \hat{c}$ et $u = \hat{d}$. Celle-ci découle des définitions de $\hat{c}$ et de $\hat{d}$ ainsi que de la relation (1.3). Ce résultat nous permet de conclure, pour $u = \hat{c}$ et $u = \hat{d}$, à l'annulation de la quantité suivante :

$$< F'(\hat{x}), p(u) > .$$

La définition de p conduit dès lors à l'assertion escomptée. Celle-ci permet, par explicitation de $\hat{c}$ et de $\hat{d}$ ainsi que par la linéarité du produit scalaire, la démonstration de l'égalité suivante :

$$< \hat{c} - \xi_+ \hat{d}, p >=< c - \xi_+ d, p > .$$

L'injection de cette dernière dans l'égalité (6.7) mène au résultat final suivant :

$$\|p\|_{\hat{x}}^2 =< c - \xi_+ d, p > .$$

Diminution constante de la fonction potentiel

Afin d'alléger les notations, nous désignons les distances $\sigma_{\hat{x}}(p)$ et $\sigma_{\hat{x}}(-p)$ respectivement par p_+ et p_-.

Théorème 6.2 *Etant donnés $\hat{x}$ l'itéré actuel et $\hat{\xi}$ la borne inférieure effective de la valeur optimale, nous déterminons ξ_+ la borne inférieure associée à la prochaine itération et p la direction de recherche comme décrit ci-dessus.*
Alors nous pouvons trouver un α strictement positif tel que

(i.) $x_+ \stackrel{def}{=} \hat{x} - \alpha p$ *appartient à* $\{x \in \text{ int } K \ : \ Bx = 0\}$

(ii.)

$$\begin{aligned} \phi(x_+; \xi_+) - \phi(\hat{x}; \hat{\xi}) &\leq -\frac{\|p\|_{\hat{x}}^2}{|p|_{\hat{x}}^2}(1 - \ln 2) \\ &\leq -(1 - \ln 2). \end{aligned}$$

Preuve:

La seconde inégalité du point $(ii.)$ découle de la propriété (3.11). Seuls le point $(i.)$ et la première inégalité du point $(ii.)$ sont donc à démontrer.

La définition de la mise à jour de $\hat{\xi}$ implique la vérification de la relation suivante :

$$\xi_+ \geq \hat{\xi}$$

et donc, par définition de la fonction potentiel primale, celle de l'inégalité ci-dessous :

$$\phi(\hat{x}; \xi_+) \leq \phi(\hat{x}; \hat{\xi}).$$

Il suffit dès lors, pour prouver la première inégalité du point $(ii.)$, de montrer l'existence d'un α strictement positif tel que

$$\Delta\phi(\alpha) \leq -\frac{\|p\|_{\hat{x}}^2}{|p|_{\hat{x}}^2}(1 - \ln 2)$$

où $\Delta\phi(\alpha) = \phi(x_+; \xi_+) - \phi(\hat{x}; \xi_+)$. Puisque nous voulons que x_+ appartienne à l'intérieur de K, nous recherchons α parmi les réels strictement inférieurs à $\frac{1}{p_+}$. Nous désignons par $\bar{c}$ la quantité $c - \xi_+ d$ et développons $\Delta\phi(\alpha)$ de la façon suivante :

$$\begin{aligned}
\Delta\phi(\alpha) &= \phi(\hat{x} - \alpha p; \xi_+) - \phi(\hat{x}; \xi_+) \\
&= \nu \ln < \bar{c}, \hat{x} - \alpha p > + F(\hat{x} - \alpha p) - \nu \ln < \bar{c}, \hat{x} > - F(\hat{x}) \\
&= \nu \ln\left(\frac{< \bar{c}, \hat{x} > - \alpha < \bar{c}, p >}{< \bar{c}, \hat{x} >}\right) + F(\hat{x} - \alpha p) - F(\hat{x}) \\
&= \nu \ln\left(1 - \alpha \frac{< \bar{c}, p >}{< \bar{c}, \hat{x} >}\right) + F(\hat{x} - \alpha p) - F(\hat{x}).
\end{aligned}$$

Nous minorons alors le rapport $\frac{<\bar{c}, p>}{<\bar{c}, \hat{x}>}$. Le point $(ii.)$ du Lemme 6.1 nous permet de majorer son dénominateur. Les deux dernières égalités établies à la fin de la subdivision précédente nous permettent, quant à elles, de réécrire son numérateur. De la sorte, nous obtenons la majoration suivante de $\Delta\phi(\alpha)$:

$$\Delta\phi(\alpha) \leq \nu \ln\left(1 - \alpha \frac{\|p\|_{\hat{x}}^2}{\nu p_-}\right) + F(\hat{x} - \alpha p) - F(\hat{x}).$$

La concavité du logarithme népérien implique la relation ci-dessous :

$$\ln\left(1 - \alpha \frac{\|p\|_{\hat{x}}^2}{\nu p_-}\right) \leq -\alpha \frac{\|p\|_{\hat{x}}^2}{\nu p_-}.$$

Dès lors, nous aboutissons à la borne suivante sur $\Delta\phi(\alpha)$:

$$\Delta\phi(\alpha) \leq -\alpha \frac{\|p\|_{\hat{x}}^2}{p_-} + F(\hat{x} - \alpha p) - F(\hat{x}).$$

Nous pouvons majorer le terme rendant compte de la variation de la barrière "ν-self-scaled" grâce au Théorème 3.2. L'application de ce théorème conduit à l'inégalité ci-dessous :

$$F(\hat{x} - \alpha p) - F(\hat{x}) \leq -\alpha < F'(\hat{x}), p > + \frac{\|p\|_{\hat{x}}^2}{p_+^2}(-\alpha p_+ - \ln(1 - \alpha p_+)),$$

satisfaite pour tout α appartenant à $[0, \frac{1}{p_+})$. Nous avons démontré à la fin de la subdivision précédente la nullité de $< F'(\hat{x}), p >$. Dès lors, la majoration de $\Delta\phi(\alpha)$ obtenue est la suivante :

$$\Delta\phi(\alpha) \leq -\alpha\frac{\|p\|_{\hat{x}}^2}{p_-} + \frac{\|p\|_{\hat{x}}^2}{p_+^2}(-\alpha p_+ - \ln(1 - \alpha p_+)).$$

Le second membre de cette inégalité peut être réécrit comme suit :

$$-\alpha\|p\|_{\hat{x}}^2(\frac{1}{p_-} + \frac{1}{p_+}) - \frac{\|p\|_{\hat{x}}^2}{p_+^2}\ln(1 - \alpha p_+).$$

On montre aisément que l'attribution de la valeur suivante :

$$\frac{1}{p_+ + p_-}$$

à α minimise cette expression. Comme il est mentionné dans les commentaires relatifs au Lemme 6.1, la non optimalité de $\hat{x}$ implique la stricte positivité de p_-. Celle-ci permet à $\alpha = \frac{1}{p_+ + p_-}$ de satisfaire la contrainte d'appartenance à $[0, \frac{1}{p_+})$ qui valide l'application antérieure du Théorème 3.2 et garantit l'appartenance de x_+ à l'intérieur de K. Par ailleurs, signalons que, puisque p est une projection sur le noyau de B relativement à $F''(\hat{x})$ et $\hat{x}$ un point strictement admissible du primal, la nullité de Bx_+ est vérifiée quelle que soit la valeur prise par α. Le point $(i.)$ est donc complètement démontré. Nous terminons la preuve de la première inégalité décrite en $(ii.)$. Dans ce but, nous remplacons α par $\frac{1}{p_+ + p_-}$ dans la majoration de $\Delta\phi(\alpha)$ obtenue. Nous aboutissons ainsi à l'inégalité ci-dessous :

$$\Delta\phi(\alpha) \leq -\frac{\|p\|_{\hat{x}}^2}{p_+ p_-} + \frac{\|p\|_{\hat{x}}^2}{p_+^2}\ln(1 + \frac{p_+}{p_-}). \tag{6.9}$$

En multipliant et divisant le second membre de cette inégalité par $\max^2\{p_+, p_-\}$, nous obtenons l'expression suivante :

$$-\frac{\|p\|_{\hat{x}}^2}{\max^2\{p_+, p_-\}}[\frac{\max^2\{p_+, p_-\}}{p_+ p_-} - \frac{\max^2\{p_+, p_-\}}{p_+^2}\ln(1 + \frac{p - +}{p_-})].$$

On démontre facilement que l'expression entre crochets vaut la quantité $g(\tau)$ définie comme suit :

$$g(\tau) \stackrel{def}{=} \frac{\max\{1, \tau^2\}}{\tau} - \frac{\max\{1, \tau^2\}}{\tau^2}\ln(1 + \tau),$$

où $\tau \stackrel{def}{=} \frac{p_+}{p_-}$. Remarquons la stricte positivité de τ. Le point $(i.)$ de la Proposition 3.2 établit la monotonie croissante de la fonction $g(\tau)$ lorsque τ est strictement supérieur à 1. D'autre part, la monotonie décroissante de cette fonction est garantie par le point $(ii.)$ de cette même proposition lorsque τ est strictement inférieur à 1. Par conséquent, l'inégalité suivante est satisfaite quel que soit τ :

$$g(\tau) \geq g(1).$$

Nous obtenons, dès lors, la majoration suivante de $\Delta\phi(\alpha)$:

$$\Delta\phi(\alpha) \leq -\frac{\|p\|_{\hat{x}}^2}{\max^2\{p_+, p_-\}} g(1).$$

Les définitions de g et de $|.|_{\hat{x}}$ permettent alors de conclure à l'inégalité attendue.

□

Rappelons que l'objectif de ce point est, étant donnés l'itéré actuel $\hat{x}$ point strictement admissible du primal et $\hat{\xi}$ borne inférieure effective de la valeur optimale, de déterminer le prochain itéré ainsi que de mettre éventuellement à jour $\hat{\xi}$ de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 6.1 et le Corollaire 6.1. Dans les deux subdivisions précédentes, nous avons explicité l'obtention d'une direction de descente $-p$ de la fonction potentiel primale ϕ en l'itéré actuel $\hat{x}$, ainsi que la mise à jour éventuelle de la borne inférieure effective $\hat{\xi}$. Il reste donc à garantir l'existence d'une longueur de pas α dans la direction $-p$ pour laquelle une diminution constante suffisante de la fonction potentiel primale ϕ et la stricte admissibilité primale de l'itéré suivant sont assurées. Le Théorème 6.2 comble ce manquement. Remarquons cependant que nous nous contentons de générer un itéré suivant x_+ qui appartient à l'intérieur de K et vérifie l'égalité $Bx_+ = 0$. De la sorte, nous omettons sciemment la contrainte de normalisation primale. Nous avons, dans l'introduction de ce point, légitimé cette omission et détaillé, sur base d'un tel itéré, l'obtention d'un itéré strictement admissible fournissant la même diminution de la fonction potentiel primale.

6.1.4 Algorithme

1. **Initialisation :** choisir x_0 un point strictement admissible du primal et ξ_0 une borne inférieure de la valeur optimale.

2. **k-ième itération** $(k \geq 0)$:

   (a) mise à jour de la borne inférieure ξ_k :

   – calcul de $\hat{c}$ et de $\hat{d}$:

$$\hat{c} \;=\; c + \frac{< c, x_k >}{\nu} F'(x_k)$$

$$\hat{d} \;=\; d + \frac{< d, x_k >}{\nu} F'(x_k);$$

– calcul de $p(\hat{c})$ et de $p(\hat{d})$: résolution du système linéaire en $(y(u), p(u))$ suivant :

$$\begin{aligned} B^* y(u) + F''(x_k) p(u) &= u \\ B p(u) &= 0 \end{aligned}$$

pour $u = \hat{c}$ et $u = \hat{d}$;

– calcul de $\tilde{x}(\xi_k)$:

$$\tilde{x}(\xi_k) = \frac{< c - \xi_k d, x_k >}{\nu} x_k + (p(\hat{c}) - \xi_k p(\hat{d}));$$

– si $\tilde{x}(\xi_k) \notin \text{int } K$: $\xi_{k+1} = \xi_k$;

– si $\tilde{x}(\xi_k) \in \text{int } K$:
calcul de $\tilde{p} = p(\hat{d}) + \frac{x_k}{\nu}$

calcul de $\sigma_{\tilde{x}(\xi_k)}(\tilde{p})$

$\xi_{k+1} = \xi_k + \frac{1}{\sigma_{\tilde{x}(\xi_k)}(\tilde{p})}$.

(b) détermination de la direction de recherche :

$$p = p(\hat{c}) - \xi_{k+1} p(\hat{d}).$$

(c) calcul de la longueur de pas :
 – calcul de $\sigma_{x_k}(p)$ et de $\sigma_{x_k}(-p)$;
 – $\alpha_k = \frac{1}{\sigma_{x_k}(p) + \sigma_{x_k}(-p)}$.

(d) détermination de l'itéré suivant :

$$x_{k+1} = \frac{x_k - \alpha_k p}{< d, x_k - \alpha_k p >}.$$

Fin de l'itération

6.1.5 Appréciation de l'extension obtenue

Notre objectif était d'étendre la méthode de réduction de potentiel primale de Karmarkar au problème primal (P). Idéalement, cette extension devait fournir un algorithme à grands pas. L'extension obtenue atteint ce double objectif. En effet, la longueur de pas utilisée ne confine pas l'itéré suivant à l'intérieur de la boule unité centrée en $\hat{x}$ et définie par la norme locale, puisque cette longueur de pas excède $\frac{1}{\|p\|_{\hat{x}}}$.

Rappelons que Yu. Nesterov et A. Nemirovskii avait déjà généralisé la méthode de Karmarkar à la programmation convexe conique mais l'algorithme qu'ils avaient obtenu était à petits pas. L'augmentation de la longueur de pas observée s'explique par la restriction aux barrières et cônes "self-scaled" ainsi que par l'introduction des nouvelles distances à la frontière de K réalisée au Chapitre 3. Ces deux facteurs valident, en effet, la borne sur la variation de la barrière F pour une longueur de pas plus grande que celle pour laquelle elle aurait été garantie lorsque seules la norme locale et une barrière self-concordante étaient disponibles. De plus, l'emploi des ces nouvelles distances permet, dans la plupart des cas, d'affiner considérablement les bornes utilisées.

Notons encore que cette dernière éventualité garantit, à chaque itération, une diminution constante de la fonction potentiel primale généralement supérieure à celle qui peut être assurée lorsque seule la norme locale est disponible. Naturellement, une accélération de la diminution de la fonction potentiel primale se traduit par un accroissement de la rapidité de convergence.

Par conséquent, la restriction aux barrières et cônes "self-scaled" et l'utilisation des nouvelles distances améliorent la rapidité de convergence de l'extension de la méthode de Karmarkar à la programmation convexe conique.

6.2 Extension de la méthode de Gonzaga

L'objectif de cette section est de fournir une méthode de réduction de potentiel primale à grands pas qui résout directement le problème primal sous la forme décrite en 5.1 et ne nécessite pas la satisfaction, par la fonction potentiel primale, de certaines propriétés. L'extension de la méthode de Karmarkar présuppose, en effet, une forme particulière des contraintes primales et l'homogénéité de degré zéro de la fonction potentiel primale. Puisque l'algorithme de Gonzaga résout le problème primal de programmation linéaire sous sa forme standard (voir [4],[8]), son extension au problème primal (P) décrit en 5.1 semble constituer une alternative qui rencontre nos exigences.

Notre généralisation de la méthode de Gonzaga utilise la fonction potentiel primale définie en (6.1), à savoir :

$$\phi(x;\xi) \stackrel{def}{=} \mu \ln(< c, x > -\xi) + F(x)$$

où μ ne peut prendre que des valeurs strictement supérieures à ν. La nécessité de cette contrainte se justifiera dans la suite.

6.2.1 Détermination des conditions de convergence vers l'optimalité primale

Proposition 6.2 *La barrière "ν-self-scaled" F considérée est bornée inférieurement sur l'ensemble $\bar{K}$ défini par*

$$\bar{K} \stackrel{def}{=} \{x \in K : \; Ax = b\tau, \; 0 \leq \tau \leq 1, \; < c, x > \leq \gamma_0\}$$

où $\gamma_0 \stackrel{def}{=} \max(< c, x_0 >, 0) + 1.$

Théorème 6.3 *Soient $\underline{F}$ le minimum de F sur $\bar{K}$, x un point strictement admissible du primal et ξ une borne inférieure de la valeur optimale supérieure ou égale à ξ_0 et telle que*

$$\phi(x; \xi) \leq \phi(x_0; \xi_0) - \Delta.$$

Alors, tant que la diminution de la fonction potentiel primale Δ satisfait la condition suivante :

$$\Delta > \mu \ln(< c, x_0 > -\xi_0) + F(x_0) - \underline{F},$$

nous avons l'inégalité ci-dessous :

$$< c, x > -\xi \leq (< c, x_0 > -\xi_0) e^{\frac{F(x_0) - \underline{F}}{\mu}} e^{\frac{-\Delta}{\mu}}.$$

Corollaire 6.2 *Supposons que nous disposions d'un point strictement admissible du primal x_0, d'une borne inférieure de la valeur optimale ξ_0 et d'un algorithme qui, à chaque itération, remplace l'itéré strictement admissible x_k et la borne inférieure de la valeur optimale ξ_k respectivement par x_{k+1} et ξ_{k+1} de même nature et ce, de telle façon que*

$$\phi(x_{k+1}; \xi_{k+1}) \leq \phi(x_k; \xi_k) - \delta$$

où δ est strictement positif.
Alors, pour un ϵ suffisamment petit (en fait, strictement inférieur à 1), nous atteignons l'ϵ-optimalité primale après K itérations, où

$$K = \lceil \frac{\mu}{\delta} (\ln(< c, x_0 > -\xi_0) + \frac{F(x_0) - \underline{F}}{\mu} + \ln(\frac{1}{\epsilon})) \rceil.$$

La Proposition 6.2, le Théorème 6.3 et le Corollaire 6.2 sont les analogues respectifs de la Proposition 6.1, du Théorème 6.1 et du Corollaire 6.1 pour l'extension de la méthode de Gonzaga. Leur preuves sont donc similaires à celles de leurs homologues. Seule celle du Théorème 6.3 diffère légèrement de celle du Théorème 6.1. En effet, cette dernière exploite l'homogénéité de degré zéro de la fonction potentiel primale associée à notre extension de la méthode de Karmarkar. Cette homogénéité n'est pas satisfaite par la fonction potentiel primale utilisée dans notre généralisation de l'algorithme de Gonzaga. La preuve du Théorème 6.3 doit dès lors avoir recours à une autre propriété de cette fonction. Celle-ci substitue l'inégalité suivante :

$$\phi(\bar{x}; \frac{\xi}{\lambda}) \leq \phi(x; \xi)$$

à l'égalité $\phi(\bar{x}, \xi) = \phi(x, \xi)$ employée dans la preuve du Théorème 6.1. Cette inégalité résulte de l'homogénéité ν-logarithmique de F et de la définition de la fonction potentiel primale utilisée. L'emploi de celle-ci ne va donc pas à l'encontre de nos objectifs.

Similairement au Théorème 6.1, le Théorème 6.3 montre qu'une diminution suffisante de la fonction potentiel primale considérée entraîne un rapprochement de l'optimalité primale. Notre extension de l'algorithme de Gonzaga doit donc s'efforcer de diminuer la valeur de sa fonction potentiel au cours de ses itérations.

Le Corollaire 6.2, au même titre que le Corollaire 6.1 pour la généralisation de la méthode de Karmarkar, établit les conditions à satisfaire pour atteindre l'ϵ-optimalité primale ainsi que le nombre d'itérations nécessaires à l'obtention de celle-ci. A nouveau, il suffit que la fonction potentiel primale décroisse à chaque itération d'une constante suffisante pour obtenir, après $O(\mu \ln(\frac{1}{\epsilon}))$ itérations, une solution primale ϵ-optimale. Soulignons, une nouvelle fois, le caractère approximatif de la borne de complexité ainsi que la souplesse des méthodes de réduction de potentiel primales qui autorisent une mise à jour de la borne inférieure.

6.2.2 Mise en oeuvre de procédures assurant la satisfaction des conditions de convergence

Supposons, comme dans le Corolloraire 6.2, qu'à chaque itération de notre généralisation de l'algorithme de Gonzaga, nous disposions d'un point strictement admissible du primal $\hat{x}$ et d'une borne inférieure de la valeur optimale $\hat{\xi}$. La question à laquelle nous répondons dans ce point est celle de la détermination de l'itéré suivant et de la mise à jour éventuelle de la borne inférieure de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 6.3 et le Corollaire 6.2.

Détermination de la direction de descente

Puisque nous voulons, à chaque itération, diminuer la valeur de la fonction potentiel primale, nous devons calculer une direction de descente de ϕ en $\hat{x}$. Nous recherchons cette direction dans le noyau de A. La présence de cette contrainte s'explique par notre volonté de générer un prochain itéré qui soit strictement admissible pour le primal. Il n'est pas nécessaire d'introduire la contrainte d'appartenance à l'intérieur de K puisque le respect de celle-ci est pris en charge par la fonction barrière F. Par ailleurs, nous préférons à la méthode de Newton une méthode de quasi-Newton. Celle-ci nous permet d'approximer en $\hat{x}$ le hessien de ϕ par celui de F qui présente l'avantage d'être défini positif. Dès lors, le problème de minimisation sous contraintes obtenu :

$$\begin{array}{ll} \min\limits_{p} & < u, p > + \frac{1}{2} < F''(\hat{x})p, p > \\ sc. & Ap = 0 \end{array}$$

où $u \stackrel{def}{=} \phi'(\hat{x}; \xi)$ et où ξ désigne une borne inférieure de la valeur optimale éventuellement mise à jour, est convexe. Nous désignons par $-p(u)$ la solution de ce problème. La convexité tant de son ensemble admissible que de sa fonction objectif nous permet de le réécrire, de façon équivalente, sous la forme d'un système de Karush-Kühn-Tücker. De la sorte, la direction de recherche $p(u)$ est obtenue en résolvant le système linéaire en $(y(u), p(u))$ suivant :

$$\begin{array}{rcl} A^*y(u) + F''(\hat{x})p(u) & = & u \\ Ap(u) & = & 0 \end{array} \tag{6.10}$$

où $y(u)$ désigne le vecteur des multiplicateurs de Lagrange associé à la contrainte d'égalité $Ap = 0$. L'existence et l'unicité de la direction de recherche $p(u)$ sont garanties par la Proposition 5.1 puisque nous constatons, à nouveau, que la direction de recherche générée s'apparente à une projection oblique. Cette constatation nous autoriserait, par ailleurs, à ne pas aborder la question de la résolution du système linéaire (6.10). Cependant, celui-ci n'est pas toujours immédiatement résolvable. Cette éventuelle impossibilité de résolution du système (6.10) provient de l'intervention de ξ dans celui-ci. En effet, u dépend de ξ. Cette borne inférieure de la valeur optimale peut avoir été mise à jour. Dans ce cas, sa valeur nous est momentanément inconnue et nous ne savons donc pas —du moins, pas complètement— résoudre le système (6.10). Dès lors, nous optons pour une méthode de résolution en deux temps dont le premier est indépendant de la valeur de ξ. Dans ce but, nous détaillons l'obtention d'une équation liant, implicitement cette fois, $p(u)$ à ξ. Les deux seules inconnues de cette équation sont $p(u)$ et une quantité λ telle que la connaissance de ξ suffit à déterminer sa valeur. Nous explicitons u et obtenons de la sorte l'égalité suivante :

$$u = \frac{\mu}{< c, \hat{x} > - \xi} + F'(\hat{x}).$$

Nous définissons ensuite les deux quantités ci-dessous :

$$\begin{aligned} \lambda &= \lambda(\xi) \stackrel{def}{=} \frac{< c, \hat{x} > -\xi}{\mu} \\ d &\stackrel{def}{=} F'(\hat{x}) \end{aligned}$$

qui permettent d'exprimer u comme suit :

$$u = \lambda^{-1}(c + \lambda d).$$

Remarquons l'indépendance de d par rapport à ξ. Comme la Proposition 5.2 établit la dépendance linéaire de la direction de recherche $p(u)$ par rapport à u, nous aboutissons à l'équation suivante :

$$p(u) = \lambda^{-1}(p(c) + \lambda p(d))$$

qui rencontre bien nos exigences. Soulignons également la positivité du coefficient λ^{-1}. La direction de recherche $p(u)$ est donc un multiple positif de $p(c) + \lambda p(d)$. Les projections obliques $p(c)$ et $p(d)$ sont définies par des systèmes linéaires indépendants de ξ et peuvent dès lors être immédiatement déterminées grâce à la méthode décrite à la fin du Chapitre 5. Ces déterminations font donc l'objet de la première phase de notre méthode de résolution du système (6.10). La seconde, quant à elle, aura lieu une fois la valeur de λ connue, c'est-à-dire après la mise à jour de la borne inférieure. Elle résoudra l'équation liant $p(u)$ à λ, $p(u)$ constituant alors l'unique inconnue de cette équation.

Mise à jour de la borne inférieure

La première partie de cette subdivision est consacrée à la description de la mise à jour de la borne inférieure. Nous soulignons également sa cohérence et sa pertinence. Finalement, nous établissons un résultat qui justifie la considération des barrières "ν-self-scaled" et des nouvelles distances introduites au Chapitre 3 dans les méthodes de réduction de potentiel primales.

Afin de déterminer la mise à jour de la borne inférieure, nous définissons, pour tout réel λ, les deux quantités suivantes :

$$\begin{aligned} \tilde{x}(\lambda) &\stackrel{def}{=} \lambda\hat{x} + p(c) + \lambda p(d) \\ \tilde{s}(\lambda) &\stackrel{def}{=} F''(\hat{x})\tilde{x}(\lambda). \end{aligned}$$

Notons que l'explicitation de $\tilde{x}(\lambda)$ et l'emploi de la première égalité décrite en (1.2) dans la définition de $\tilde{s}(\lambda)$, nous permettent d'exprimer cette dernière quantité comme suit :

$$\tilde{s}(\lambda) = F''(\hat{x})(p(c) + \lambda p(d)) - \lambda F'(\hat{x}). \tag{6.11}$$

En combinant les secondes équations des systèmes définissant $p(c)$ et $p(d)$, et en utilisant l'égalité (6.11) ainsi que la définition de d, nous aboutissons à l'égalité suivante :

$$A^*(y(c) + \lambda y(d)) + \tilde{s}(\lambda) = c.$$

Par conséquent, tant que $\tilde{s}(\lambda)$ appartient au cône dual K^*, le couple $(y(c) + \lambda y(d), \tilde{s}(\lambda))$ est un point admissible du dual (D). La fonction objectif duale évaluée en ce point vaut, par la définition du saut de dualité et la stricte admissibilité primale de $\hat{x}$, $< c, \hat{x} > - < \tilde{s}(\lambda), \hat{x} >$. Cette valeur constitue, lorsque l'appartenance de $\tilde{s}(\lambda)$ à K^* est satisfaite, une borne inférieure de la valeur optimale commune à (P) et à (D). Nous définissons alors la quantité suivante :

$$\hat{\lambda} \stackrel{def}{=} \begin{cases} \min\{\lambda \in I\!R : \tilde{s}(\lambda) \in K^*\} & \text{si } \{\lambda \in I\!R : \tilde{s}(\lambda) \in K^*\} \text{ est non vide,} \\ +\infty & \text{sinon} \end{cases}$$

et attribuons à ξ_+, la borne inférieure associée à l'itération suivante, la valeur ci-dessous :

$$\max(\hat{\xi}, < c, \hat{x} > - < \tilde{s}(\hat{\lambda}), \hat{x} >).$$

Nous constatons que la valeur du second argument de ce maximum n'est définie que pour des valeurs réelles de $\hat{\lambda}$. Nous lui affectons la valeur $-\infty$ lorsque $\hat{\lambda}$ vaut $+\infty$. Le cas $\hat{\lambda}$ égale $-\infty$ n'est pas à envisager si, comme nous l'avons implicitement supposé, la borne inférieure de $\{\lambda \in I\!R : \tilde{s}(\lambda) \in K^*\}$ est atteinte. Cette assertion fait l'objet de la proposition suivante.

Proposition 6.3 *La borne inférieure de l'ensemble suivant :*

$$\{\lambda \in I\!R : \tilde{s}(\lambda) \in K^*\}$$

est atteinte, si celui-ci est non vide.

Preuve:

1. Nous montrons que, si nous excluons le cas où b est nul, le saut de dualité $< \tilde{s}(\lambda), \hat{x} >$ est une fonction strictement croissante en λ . Dans ce but, nous exprimons tout d'abord le produit scalaire $< \tilde{s}(\lambda), \hat{x} >$ comme une fonction affine en λ. La définition de $\tilde{s}(\lambda)$, l'hypothèse de symétrie imposée aux hessiens de F, la première égalité décrite en (1.2) et la définition de la norme $\|.\|_{\hat{x}}$ légitiment l'égalité suivante :

$$< \tilde{s}(\lambda), \hat{x} > = < -F'(\hat{x}), p(c) > + \lambda[\|\hat{x}\|_{\hat{x}}^2 - < F'(\hat{x}), p(d) >].$$

La première égalité décrite en (1.2) permet également la vérification de la relation ci-dessous:

$$\|\hat{x}\|_{\hat{x}} = \|F'(\hat{x})\|^*_{\hat{x}}.$$

Par ailleurs, les définitions de d et de $p(d)$ ainsi que la Proposition 5.3 expliquent l'égalité suivante:

$$< F'(\hat{x}), p(d) >= \|p(d)\|^2_{\hat{x}}.$$

Dès lors, nous obtenons l'expression suivante du saut de dualité considéré:

$$< \tilde{s}(\lambda), \hat{x} >=< -F'(\hat{x}), p(c) > +\lambda[(\|F'(\hat{x})\|^*_{\hat{x}})^2 - \|p(d)\|^2_{\hat{x}}].$$

Nous montrons ensuite que le second membre de l'égalité ci-dessus est une fonction strictement croissante en λ. Nous devons donc montrer l'inégalité suivante:

$$(\|F'(\hat{x})\|^*_{\hat{x}})^2 > \|p(d)\|^2_{\hat{x}}.$$

La Proposition 5.4 permet une majoration stricte de $\|p(d)\|^2_{\hat{x}}$. En effet, ce résultat établit, pour tout y appartenant à Y, l'inégalité suivante:

$$\|p(d)\|^2_{\hat{x}} \leq (\|d - A^*y\|^*_{\hat{x}})^2$$

qui devient une égalité seulement pour y valant $y(d)$, l'unique minimum de $\|d - A^*y\|^*_{\hat{x}}$. Puisque l'opérateur A^* est défini sur Y, son noyau est contenu dans ce dernier. Cette inclusion et la définition de d impliquent qu'il suffit, pour montrer la stricte croissance du saut de dualité, de trouver un élément y du noyau de A^* qui soit distinct de $y(d)$. La linéarité et l'injectivité de l'opérateur A^* réduisent son noyau à l'élément nul. Par conséquent, nous pouvons nous contenter de montrer que $y(d)$ est différent de zéro. $y(d)$ solutionne le système linéaire en $(y(d), p(d))$ suivant:

$$\begin{aligned} Ap(d) &= 0 \\ A^*y(d) + F''(\hat{x})p(d) &= d \end{aligned}$$

Supposons, par l'absurde, $y(d)$ nul. Ce système devient dès lors, grâce également à la définition de d, le suivant:

$$\begin{aligned} Ap(d) &= 0 \\ F''(\hat{x})p(d) &= F'(\hat{x}) \end{aligned} \tag{6.12}$$

La première égalité décrite en (1.2) implique que $-\hat{x}$ solutionne la seconde équation de (6.12). Or celle-ci admet une solution unique étant donné le caractère défini positif imposé aux hessiens de F. Dès lors, nous obtenons l'égalité suivante:

$$p(d) = -\hat{x}$$

qui conduit à une contradiction si nous supposons b non nul. En effet, $-\hat{x}$ n'appartient pas au noyau de A puisqu'il satisfait à $A\hat{x} = b$ et ne peut donc satisfaire la première équation de (6.12).

2. Nous justifions l'exclusion du cas b égale 0. Sous cette hypothèse, nous connaissons effectivement l'ensemble des solutions primales : celui-ci se réduit à l'élément nul. En effet, supposons, par l'absurde, l'existence d'une autre solution primale x^*. Deux cas de figure sont possibles : soit la fonction objectif primale évaluée en x^* égale son évaluation en zéro, autrement dit est nulle; soit celle-ci est strictement négative. Dans le premier cas, tout multiple positif de x^* est admissible et optimal. Dès lors, l'ensemble des solutions primales est non borné, ce qui contredit nos hypothèses. Dans le second cas, comme tout multiple positif de x^* est admissible, la valeur de la fonction objectif primale n'est pas bornée inférieurement. Cette constatation conduit, à nouveau, à une contradiction puisque nos hypothèses garantissent l'existence de solutions primales.

3. Nous montrons que la borne inférieure de $\{\lambda \in I\!R \ : \ \tilde{s}(\lambda) \in K^*\}$ est atteinte. Dans ce but, nous modifions momentanément la définition de $\hat{\lambda}$: cette quantité désignera, dans la suite de cette preuve, la borne inférieure de l'ensemble $\{\lambda \in I\!R \ : \ \tilde{s}(\lambda) \in K^*\}$. Nous supposons alors , par l'absurde, que $\hat{\lambda}$ égale $-\infty$. Cette supposition et la définition de $\hat{\lambda}$ impliquent, quel que soit λ réel, l'appartenance suivante :

$$\tilde{s}(\lambda) \in K^*$$

et donc, par définition de K^*, la satisfaction par tout λ réel de l'inégalité ci-dessous :

$$< \tilde{s}(\lambda), \hat{x} > \geq 0. \tag{6.13}$$

Nous avons en premier lieu exprimer le saut de dualité comme suit :

$$< \tilde{s}(\lambda), \hat{x} > = < -F'(\hat{x}), p(c) > + \lambda[(\|F'(\hat{x})\|_{\hat{x}}^*)^2 - \|p(d)\|_{\hat{x}}^2]$$

et montrer qu'il s'agit d'une fonction strictement croissante en λ. Dès lors, si nous définissons la fonction g par :

$$g(\lambda) \overset{def}{=} \lambda[(\|F'(\hat{x})\|_{\hat{x}}^*)^2 - \|p(d)\|_{\hat{x}}^2],$$

nous obtenons d'une part, que sa dérivée première est strictement positive et d'autre part, l'égalité suivante :

$$< \tilde{s}(\lambda), \hat{x} > = < -F'(\hat{x}), p(c) > + \lambda \frac{d}{d\lambda} g(\lambda).$$

Nous faisons alors tendre λ vers $\hat{\lambda}$, c'est-à-dire vers $-\infty$ et obtenons, de la sorte, une contradiction. En effet, le saut de dualité tend ainsi vers $-\infty$ alors que nous avons établi la constance de sa positivité en (6.13).

□

La Proposition 6.3 établit la justesse de la définition initiale de $\hat{\lambda}$ et par suite, la cohérence de notre stratégie de mise à jour de la borne inférieure. Le but de la mise à jour de la borne inférieure est de fournir, lorsque cela est possible, une borne inférieure de la valeur optimale plus fine que celle utilisée à l'itération précédente. La mise à jour décrite ci-dessus atteint cet objectif. En effet, le premier argument du maximum la définissant garantit que la nouvelle borne inférieure est au moins aussi fine que la précédente. D'autre part, le second argument fournit la meilleure borne inférieure de la valeur optimale de la forme suivante :

$$< c, \hat{x} > - < \tilde{s}(\lambda), \hat{x} > .$$

La définition de $\hat{\lambda}$ et la stricte croissance du saut de dualité établie dans la preuve de la Proposition 6.3 entraînent, effectivement, que $< c, \hat{x} > - < \tilde{s}(\lambda), \hat{x} >$, considéré comme une fonction de λ définie sur l'ensemble $\{\lambda \in I\!R \; : \; \tilde{s}(\lambda) \in K^*\}$ supposé non vide, est maximal en $\hat{\lambda}$.

Par ailleurs, nous pouvons exprimer la direction de recherche $p(u)$ comme un multiple positif de p, p étant défini par :

$$p \stackrel{def}{=} p(c) + \lambda_+ p(d)$$

où $\lambda_+ \stackrel{def}{=} \lambda(\xi_+)$.

Lemme 6.2

(i.) Si $\mu \geq \nu + \sqrt{\nu}$, nous obtenons la majoration suivante de λ_+ :

$$\lambda_+ \leq \|p\|_{\hat{x}}.$$

(ii.) Si $\mu \geq 2\nu$, nous obtenons la majoration suivante de λ_+ :

$$\lambda_+ \leq |p|_{\hat{x}}.$$

Preuve:
Ce résultat se démontre par l'absurde. La preuve du point $(i.)$ suppose donc satisfaite l'inégalité suivante :

$$\lambda_+ > \|p\|_{\hat{x}}$$

et celle du point $(ii.)$ admet la minoration stricte suivante de λ_+ :

$$\lambda_+ > |p|_{\hat{x}}.$$

Dans les deux cas, nous avons que :

$$\lambda_+ > \sigma_{\hat{x}}(-p).$$

Cette conséquence est déduite, pour le point $(i.)$, des propriétés décrites en (3.11) et, pour le point $(ii.)$, de la définition de $|.|_{\hat{x}}$. Elle entraîne l'appartenance de $\lambda_+\hat{x}+p$ à K. D'autre part, les définitions de $\tilde{s}(\lambda_+)$, de $\tilde{x}(\lambda_+)$ et de p permettent d'exprimer $\tilde{s}(\lambda_+)$ comme suit :

$$\tilde{s}(\lambda_+) = F''(\hat{x})(\lambda_+\hat{x} + p). \tag{6.14}$$

Dès lors, la stricte admissibilité primale de $\hat{x}$, l'appartenance de $\lambda_+\hat{x} + p$ à K déduite ci-dessus et le point $(iii.)$ du Théorème 2.1 garantissent l'appartenance de $\tilde{s}(\lambda_+)$ au cône dual K^*. Par conséquent, $< c, \hat{x} > - < \tilde{s}(\lambda_+), \hat{x} >$ est une borne inférieure de la valeur optimale. Puisque ξ_+ a été définie comme la meilleure borne inférieure de la valeur optimale de la forme $< c, \hat{x} > - < \tilde{s}(\lambda), \hat{x} >$, nous obtenons l'inégalité suivante :

$$< c, \hat{x} > - < \tilde{s}(\lambda_+), \hat{x} > \leq \xi_+$$

ou, de façon équivalente,

$$< c, \hat{x} > -\xi_+ \leq < \tilde{s}(\lambda_+), \hat{x} > .$$

Nous développons alors le second membre de l'inégalité ci-dessus. La relation (6.14), la première égalité décrite en (1.4) et l'hypothèse de symétrie imposée aux hessiens de F combinée à la première propriété décrite en (1.2) valident l'égalité suivante :

$$< \tilde{s}(\lambda_+), \hat{x} >= \lambda_+\nu + < -F'(\hat{x}), p > .$$

Par conséquent, nous pouvons majorer $< c, \hat{x} > -\xi_+$ de la façon suivante :

$$< c, \hat{x} > -\xi_+ \leq \lambda_+\nu + < -F'(\hat{x}), p > . \tag{6.15}$$

Nous démontrons, par l'absurde, le point $(i.)$. L'inégalité (3.9) nous permet de majorer le second membre de (6.15). Nous obtenons de la sorte la relation ci-dessous :

$$< c, \hat{x} > -\xi_+ \leq \lambda_+\nu + \sqrt{\nu}\|p\|_{\hat{x}}.$$

L'utilisation de l'hypothèse par l'absurde $\lambda_+ > \|p\|_{\hat{x}}$ et de la supposition inhérente au point $(i.)$ $\mu \geq \nu + \sqrt{\nu}$, nous conduit à l'inégalité stricte suivante :

$$< c, \hat{x} > -\xi_+ < \mu\lambda_+.$$

Celle-ci est équivalente à la relation ci-dessous :

$$\lambda_+ > \frac{< c, \hat{x} > -\xi_+}{\mu},$$

qui contredit la définition de λ_+.
La preuve du point (*ii.*) utilise le même schéma de démonstration. Il suffit d'employer l'inégalité (3.10) au lieu de celle décrite en (3.9) et de remarquer que l'hypothèse par l'absurde, $\lambda_+ > |p|_{\hat{x}}$, implique l'inégalité suivante :

$$\lambda_+ > \sigma_{\hat{x}}(p).$$

$\square$

Le Lemme 6.2 est l'analogue du point (*ii.*) du Lemme 6.1 pour l'extension de l'algorithme de Gonzaga. Il fournit des bornes qui permettront la démonstration d'un théorème qui garantit, à chaque itération, une diminution constante de la fonction potentiel primale. L'obtention d'une telle borne dans la généralisation de la méthode de Karmarkar, ne nécessitait pas l'exploitation du caractère "ν-self-scaled" de la barrière. La preuve du Lemme 6.2 y a, quant à elle, recours. De plus, malgré la puissance des outils utilisés, la borne dégagée dans le cas $\mu \geq \nu + \sqrt{\nu}$ n'est pas plus fine que celle qui aurait été établie si la norme locale définie par le hessien de la barrière était la seule mesure de la distance à la frontière de K disponible. Cette constatation compromet sérieusement nos chances d'obtention d'une grande diminution de la fonction potentiel primale à chaque itération. Toutefois, un renforcement de la contrainte imposée à μ permet d'établir une borne de finesse comparable à celle obtenue dans le point (*ii.*) du Lemme 6.1. Par conséquent, une grande diminution de la fonction potentiel primale semble accessible, à chaque itération, pour des valeurs de μ excédant 2ν.

Diminution constante de la fonction potentiel

Afin d'alléger les notations, nous désignons les distances $\sigma_{\hat{x}}(p)$ et $\sigma_{\hat{x}}(-p)$ respectivement par p_+ et p_-.

Théorème 6.4 *Etant donnés $\hat{x}$ l'itéré actuel et $\hat{\xi}$ la borne inférieure effective de la valeur optimale, nous déterminons ξ_+ la borne inférieure associée à la prochaine itération et p la direction de recherche comme décrit ci-dessus.*
Si $\mu \geq \nu + \sqrt{\nu}$, nous pouvons trouver un α strictement positif tel que

(i.) $x_+ \stackrel{def}{=} \hat{x} - \alpha p$ *est un point strictement admissible du primal*

(ii.)

$$\begin{aligned} \phi(x_+;\xi_+) - \phi(\hat{x};\hat{\xi}) &\leq -\frac{\|p\|_{\hat{x}}^2}{\max^2\{p_+, \|p\|_{\hat{x}}^2\}}(1-\ln 2) \\ &= -(1-\ln 2) \end{aligned}$$

Si $\mu \geq 2\nu$, nous pouvons trouver un α strictement positif tel que

(i.) $x_+ \stackrel{def}{=} \hat{x} - \alpha p$ *est un point strictement admissible du primal*

(ii.)

$$\begin{aligned}
\phi(x_+;\xi_+) - \phi(\hat{x};\hat{\xi}) &\leq -\frac{\|p\|_{\hat{x}}^2}{\max^2\{p_+, |p|_{\hat{x}}\}}(1-\ln 2) \\
&= -\frac{\|p\|_{\hat{x}}^2}{|p|_{\hat{x}}^2}(1-\ln 2) \\
&\leq -(1-\ln 2)
\end{aligned}$$

Preuve:
Nous nous intéressons, tout d'abord, à l'appartenance de x_+ à l'ensemble primal strictement admissible. La linéarité des projections obliques établie dans la Proposition 5.2 implique que p est la projection de $c + \lambda_+ d$ sur le noyau de A relativement à $F''(\hat{x})$. Dès lors, x_+ vérifie la contrainte $Ax_+ = b$ quelle que soit la valeur prise par α. La contrainte d'appartenance de x_+ à l'intérieur de K est, quant à elle, satisfaite lorsque α appartient à l'intervalle $[0, \frac{1}{p_+})$. Nous recherchons donc α parmi les réels strictement inférieurs à $\frac{1}{p_+}$.
Nous détaillons ensuite une majoration de la différence $\phi(x_+;\xi_+) - \phi(\hat{x};\hat{\xi})$ commune aux deux cas envisagés. La définition de la mise à jour implique la vérification de la relation suivante:

$$\xi_+ \geq \hat{\xi}$$

et donc, par définition de la fonction potentiel primale, celle de l'inégalité ci-dessous:

$$\phi(\hat{x};\xi_+) \leq \phi(\hat{x};\hat{\xi}).$$

Par conséquent, toute majoration de $\Delta\phi(\alpha)$ où $\Delta\phi(\alpha) \stackrel{def}{=} \phi(x_+;\xi_+) - \phi(\hat{x};\xi_+)$, reste valable pour $\phi(x_+;\xi_+) - \phi(\hat{x};\hat{\xi})$. Nous développons $\Delta\phi(\alpha)$ de la façon suivante:

$$\begin{aligned}
\Delta\phi(\alpha) &= \phi(\hat{x} - \alpha p;\xi_+) - \phi(\hat{x};\xi_+) \\
&= \mu \ln(< c, \hat{x} - \alpha p > -\xi_+) + F(\hat{x} - \alpha p) - \mu \ln(< c, \hat{x} > -\xi_+) - F(\hat{x}) \\
&= \mu \ln(\frac{< c, \hat{x} > -\alpha < c, p > -\xi_+}{< c, \hat{x} > -\xi_+}) + F(\hat{x} - \alpha p) - F(\hat{x}) \\
&= \mu \ln(1 - \frac{\alpha < c, p >}{< c, \hat{x} > -\xi_+}) + F(\hat{x} - \alpha p) - F(\hat{x})
\end{aligned}$$

La concavité du logarithme népérien implique la relation ci-dessous:

$$\ln(1 - \frac{\alpha < c, p >}{< c, \hat{x} > -\xi_+}) \leq -\frac{\alpha < c, p >}{< c, \hat{x} > -\xi_+}.$$

D'autre part, le Théorème 3.2 permet, pour tout α appartenant à $[0, \frac{1}{p_+})$, de majorer le terme rendant compte de la variation de la barrière "ν-self-scaled" comme suit :

$$F(\hat{x} - \alpha p) - F(\hat{x}) \leq -\alpha < F'(x), p > + \frac{\|p\|_{\hat{x}}^2}{p_+^2}(-\alpha p_+ - \ln(1 - \alpha p_+)).$$

Dès lors, nous pouvons borner $\Delta\phi(\alpha)$ de la façon suivante :

$$\Delta\phi(\alpha) \leq -\alpha < c, p > \frac{\mu}{< c, \hat{x} > -\xi_+} - \alpha < F'(x), p > + \frac{\|p\|_{\hat{x}}^2}{p_+^2}(-\alpha p_+ - \ln(1 - \alpha p_+)).$$

Le second membre de cette inégalité peut, grâce à la définition de λ_+ et à une mise en évidence, être réécrit comme suit :

$$\frac{-\alpha}{\lambda_+} < c + \lambda_+ F'(\hat{x}), p > + \frac{\|p\|_{\hat{x}}^2}{p_+^2}(-\alpha p_+ - \ln(1 - \alpha p_+)).$$

Par ailleurs, nous avons l'égalité suivante :

$$\|p\|_{\hat{x}}^2 = < c + \lambda_+ F'(\hat{x}), p > .$$

En effet, nous obtenons, par les définitions de la norme $\|.\|_{\hat{x}}$ et de p, la relation suivante :

$$\|p\|_{\hat{x}}^2 = < F''(\hat{x})p(c), p > + \lambda_+ < F''(\hat{x})p(d), p > .$$

Les secondes équations des systèmes définissant $p(c)$ et $p(d)$ impliquent l'égalité ci-dessous :

$$F''(\hat{x})p(u) = u - A^* y(u)$$

pour $u = c$ et $u = d$. Dès lors, nous pouvons réécrire $\|p\|_{\hat{x}}^2$ comme suit :

$$\|p\|_{\hat{x}}^2 = < c, p > - < A^* y(c), p > + \lambda_+ < d, p > - \lambda_+ < A^* y(d), p > .$$

Le second membre de cette équation se réduit à

$$< c + \lambda_+ d, p >$$

grâce aux premières équations des systèmes définissant $p(c)$ et $p(d)$ ainsi qu'à la définition de l'adjointe de A. La définition de d nous permet alors d'aboutir à l'égalité escomptée. Nous injectons celle-ci dans notre majoration de $\Delta\phi(\alpha)$ obtenue et aboutissons, de la sorte, à l'inégalité suivante :

$$\Delta\phi(\alpha) \leq -\alpha \|p\|_{\hat{x}}^2 (\frac{1}{\lambda_+} + \frac{1}{p_+}) - \frac{\|p\|_{\hat{x}}^2}{p_+^2} \ln(1 - \alpha p_+). \tag{6.16}$$

Nous traitons, en premier lieu, le cas où $\mu \geq \nu + \sqrt{\nu}$. Le Lemme 6.2 permet une majoration de λ_+. Dès lors, l'inégalité (6.16) devient la suivante :

$$\Delta\phi(\alpha) \leq -\alpha \|p\|_{\hat{x}}^2 (\frac{1}{\|p\|_{\hat{x}}} + \frac{1}{p_+}) - \frac{\|p\|_{\hat{x}}^2}{p_+^2} \ln(1 - \alpha p_+).$$

On remarque que l'attribution de la valeur ci-dessous :

$$\frac{1}{p_+ + \|p\|_{\hat{x}}}$$

à α minimise le second membre de cette inégalité. Soulignons également la stricte positivité de $\|p\|_{\hat{x}}$. Celle-ci permet à $\alpha = \frac{1}{p_+ + \|p\|_{\hat{x}}}$ de satisfaire la contrainte d'appartenance à $[0, \frac{1}{p_+})$ qui valide l'application antérieure du Théorème 3.2 et garantit l'appartenance de x_+ à l'intérieur de K. Le remplacement de α par $\frac{1}{p_+ + \|p\|_{\hat{x}}}$ dans la majoration de $\Delta\phi(\alpha)$ obtenue, nous conduit à l'inégalité suivante :

$$\Delta\phi(\alpha) \leq \frac{-\|p\|_{\hat{x}}^2}{p_+\|p\|_{\hat{x}}} + \frac{\|p\|_{\hat{x}}^2}{p_+^2}\ln(1 + \frac{p_+}{\|p\|_{\hat{x}}})$$

qui coïncide avec celle décrite en (6.9), à la substitution de $\|p\|_{\hat{x}}$ à p_- près. Par conséquent, un raisonnement similaire à celui détaillé dans la preuve du Théorème 6.2 nous conduit au résultat suivant :

$$\Delta\phi(\alpha) \leq \frac{-\|p\|_{\hat{x}}^2}{\max^2\{p_+, \|p\|_{\hat{x}}\}}(1 - \ln 2)$$

qui conclut à la satisfaction de la première inégalité décrite dans le point $(ii.)$. L'égalité décrite dans ce même point découle d'une des propriétés répertoriées en (3.11). Nous démontrons, finalement, le point $(ii.)$ du cas $\mu \geq 2\nu$. A nouveau, le Lemme 6.2 permet une majoration du λ_+ intervenant dans l'inégalité (6.16) qui devient dès lors la suivante :

$$\Delta\phi(\alpha) \leq -\alpha\|p\|_{\hat{x}}^2(\frac{1}{|p|_{\hat{x}}} + \frac{1}{p_+}) - \frac{\|p\|_{\hat{x}}^2}{p_+^2}\ln(1 - \alpha p_+).$$

L'attribution de la valeur ci-dessous :

$$\frac{1}{p_+ + |p|_{\hat{x}}}$$

à α minimise le second membre de cette inégalité. Remarquons la stricte positivité de $|p|_{\hat{x}}$. Cette constatation valide à nouveau l'application antérieure du Théorème 3.2 et garantit l'appartenance de x_+ à l'intérieur de K. Le remplacement de α par $\frac{1}{p_+ + |p|_{\hat{x}}}$ dans la majoration de $\Delta\phi(\alpha)$ obtenue conduit à une inégalité qui coïncide avec celle décrite en (6.9), à la substitution de $|p|_{\hat{x}}$ à p_- près. Dès lors, un raisonnement similaire à celui détaillé dans la preuve du Théorème 6.2 mène à l'inégalité suivante :

$$\Delta\phi(\alpha) \leq \frac{-\|p\|_{\hat{x}}^2}{\max^2\{p_+, |p|_{\hat{x}}\}}(1 - \ln 2)$$

qui implique la satisfaction de la première inégalité du point $(ii.)$. L'égalité décrite dans ce même point est une conséquence directe des définitions de p_+ et de $|p|_{\hat{x}}$. Quant à la seconde inégalité de ce point, elle découle des propriétés décrites en (3.11).

□

Rappelons que l'objectif de ce point est, étant donnés l'itéré actuel $\hat{x}$ point strictement admissible du primal et $\hat{\xi}$ borne inférieure effective de la valeur optimale, de déterminer le prochain itéré ainsi que de mettre éventuellement à jour $\hat{\xi}$ de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 6.3 et le Corollaire 6.2. Dans les deux subdivisions précédentes, nous avons explicité l'obtention d'une direction de descente $-p$ de la fonction potentiel primale ϕ en l'itéré actuel $\hat{x}$, ainsi que la mise à jour éventuelle de la borne inférieure effective $\hat{\xi}$. Il reste donc à garantir l'existence d'une longueur de pas α dans la direction $-p$ pour laquelle une diminution constante suffisante de la fonction potentiel primale ϕ et la stricte admissibilité primale du prochain itéré sont assurées. Le Théorème 6.4 comble ce manquement.

6.2.3 Algorithme

1. **Initialisation :** choisir x_0 un point strictement admissible du primal et ξ_0 une borne inférieure de la valeur optimale.

2. **k-ième itération** $(k \geq 0)$ **:**

   (a) mise à jour de la borne inférieure ξ_k :

   – calcul de d :
   $$d = F'(x_k);$$

   – calcul de $p(c)$ et de $p(d)$: résolution du système linéaire en $(y(u), p(u))$ suivant :
   $$\begin{aligned} A^*y(u) + F''(x_k)p(u) &= u \\ Ap(u) &= 0 \end{aligned}$$
   pour $u = c$ et $u = d$;

   – calcul de $\hat{\lambda}_k$:
   $$\hat{\lambda}_k \stackrel{def}{=} \begin{cases} \min\{\lambda \in I\!R \ : \ \tilde{s}(\lambda) \in K^*\} & \text{si } \{\lambda \in I\!R \ : \ \tilde{s}(\lambda) \in K^*\} \text{ est non vide,} \\ +\infty & \text{sinon} \end{cases}$$
   où $\tilde{s}(\lambda) = F''(x_k)[\lambda x_k + p(c) + \lambda p(d)]$;

   – calcul de ξ_{k+1} et de λ_{k+1} :
   $$\begin{aligned} \xi_{k+1} &= \max\{\xi_k, < c, x_k > - < \tilde{s}(\hat{\lambda}_k), x_k >\} \\ & \quad \text{[le second argument du maximum vaut } -\infty \text{ lorsque } \hat{\lambda}_k \text{ égale } +\infty] \\ \lambda_{k+1} &= \lambda(\xi_{k+1}) = \frac{< c, x_k > - \xi_{k+1}}{\mu}. \end{aligned}$$

(b) détermination de la direction de recherche :

$$p = p(c) + \lambda_{k+1} p(d).$$

(c) calcul de la longueur de pas :

– calcul de $\sigma_{x_k}(p)$ et de $\begin{cases} \|p\|_{x_k} & \text{si } \mu \geq \nu + \sqrt{\nu}, \\ |p|_{x_k} & \text{si } \mu \geq 2\nu; \end{cases}$

– $\alpha_k = \begin{cases} \frac{1}{\sigma_{x_k}(p) + \|p\|_{x_k}} & \text{si } \mu \geq \nu + \sqrt{\nu}, \\ \frac{1}{\sigma_{x_k}(p) + |p|_{x_k}} & \text{si } \mu \geq 2\nu. \end{cases}$

(d) détermination de l'itéré suivant :

$$x_{k+1} = x_k - \alpha_k p.$$

Fin de l'itération

6.2.4 Appréciation de l'extension obtenue

Notre objectif était d'étendre la méthode de réduction de potentiel primale de Gonzaga au problème primal (P). Idéalement, cette extension devait fournir un algorithme à grands pas. L'extension obtenue dans le cas $\mu \geq \nu + \sqrt{\nu}$ atteint uniquement l'objectif de base. En effet, la longueur de pas utilisée dans ce cas confine l'itéré suivant à l'intérieur de la boule unité centrée en $\hat{x}$ et définie par la norme locale, puisque cette longueur de pas est bornée par $\frac{1}{\|p\|_{\hat{x}}}$. De plus, seule une petite diminution de la fonction potentiel primale est garantie à chaque itération. Toutefois, en renforçant la contrainte imposée à μ, nous sommes parvenus à une extension à grands pas. Une grande diminution de la fonction potentiel primale est, par ailleurs, accessible à chacune des itérations de celle-ci. Signalons que, dans la généralisation de la méthode de Karmarkar, un algorithme à grands pas et une grande diminution de la fonction potentiel primale étaient déjà obtenus pour μ valant ν.

Par conséquent, la forme particulière des contraintes primales et l'homogénéité de degré zéro de la fonction potentiel primale inhérentes à notre extension de l'algorithme de Karmarkar, ne sont pas nécessaires à l'élaboration d'une méthode de réduction de potentiel primale qui converge. Cependant, notre volonté d'obtention d'une méthode à grands pas nécessite l'introduction d'une contrainte supplémentaire sur μ. Les valeurs attribuées à ce paramètre doivent, en effet, excéder 2μ.

Chapitre 7

Méthode de réduction de potentiel primale-duale symétrique

Ce chapitre est consacré au développement d'une méthode de réduction de potentiel qui résout simultanément et symétriquement les problèmes primal (P) et dual (D) décrits en 5.1. Puisque l'algorithme de Kojima, Mizuno et Yoshise implémente une telle méthode pour les problèmes primal et dual de programmation linéaire (voir [5], [8]), son extension aux problèmes (P) et (D) semble constituer une alternative qui rencontre nos exigences.

Notre généralisation de la méthode de Kojima, Mizuno et Yoshise utilise la fonction potentiel primale-duale définie par

$$\phi(x,s) \stackrel{def}{=} (\nu+\rho)\ln < s,x > +F(x) + F_*(s)$$

où ρ désigne une constante supérieure ou égale à $\sqrt{\nu}$ et x un point strictement admissible du primal. Quant à s, il provient d'un point strictement admissible du dual (y,s). D'autre part, nous supposons satisfaites les hypothèses répertoriées en 5.1.

Ce chapitre est structuré comme suit. Nous dégageons, tout d'abord, les conditions que doit satisfaire notre extension pour converger vers l'optimalité primale-duale. Nous détaillons, ensuite, l'obtention de celle-ci en élaborant une procédure et en établissant un résultat qui garantissent la satisfaction de ces conditions de convergence. Finalement, nous apprécions l'algorithme ainsi obtenu.

7.1 Détermination des conditions de convergence vers l'optimalité primale-duale

Théorème 7.1 *Soient x et (y,s) des points strictement admissibles respectivement du primal et du dual tels que*

$$\phi(x,s) \leq \phi(x_0,s_0) - \Delta.$$

Alors, le saut de dualité correspondant à (x,s) vérifie l'inégalité ci-dessous :

$$< s, x > \leq < s_0, x_0 > R(x_0,s_0) e^{\frac{-\Delta}{\rho}}$$

où $R(x_0,s_0) \stackrel{def}{=} e^{\frac{\nu \ln < s_0, x_0 > + F(x_0) + F_*(s_0) - \nu \ln \nu + \nu}{\rho}}$.

Preuve:

Yu. Nesterov et A. Nemirovskii ont établi dans [2, Proposition 4.5.1] le résultat suivant :

Proposition 1 *Soient γ un réel strictement positif, F une barrière ν-normale pour le cône primal K, (x,s) un couple primal-dual strictement admissible et (x_0, s_0) le couple primal-dual strictement admissible initial.*

Alors, le rapport des sauts de dualité définis respectivement par (x,s) et par (x_0, s_0) vérifie l'inégalité ci-dessous :

$$\frac{< s, x >}{< s_0, x_0 >} \leq R(x_0,s_0) e^{\frac{V_*(x,s) - V_*(x_0,s_0)}{\gamma\sqrt{\nu}}}$$

où

$$\begin{aligned} R(x_0,s_0) &\stackrel{def}{=} e^{\frac{\nu \ln < s_0, x_0 > + F(x_0) + F_*(s_0) - \nu \ln \nu + \nu}{\gamma\sqrt{\nu}}} \\ V_*(x,s) &\stackrel{def}{=} (\nu + \gamma\sqrt{\nu}) \ln < s, x > + F(x) + F_*(s). \end{aligned}$$

Nous appliquons ce résultat avec l'identification $\gamma = \frac{\rho}{\sqrt{\nu}}$. Cette identification fait coïncider $V_*(x,s)$ avec $\phi(x,s)$. Nous obtenons, dès lors, par application de la Proposition 1, l'inégalité suivante :

$$\frac{< s, x >}{< s_0, x_0 >} \leq R(x_0,s_0) e^{\frac{\phi(x,s) - \phi(x_0,s_0)}{\rho}}$$

où

$$R(x_0,s_0) \stackrel{def}{=} e^{\frac{\nu \ln < s_0, x_0 > + F(x_0) + F_*(s_0) - \nu \ln \nu + \nu}{\rho}}.$$

L'exponentielle intervenant explicitement dans l'inégalité peut, grâce l'hypothèse relative à la diminution de la fonction potentiel primale-duale, à la stricte positivité de ρ ainsi qu'à la croissance de l'exponentielle, être majorée par

$$e^{\frac{-\Delta}{\rho}}.$$

D'autre part, la non-optimalité de (x_0, y_0, s_0) entraîne la stricte positivité du saut de dualité correspondant à (x_0, s_0). Par conséquent, nous aboutissons bien ainsi à l'inégalité attendue, à savoir :

$$< s, x > \leq < s_0, x_0 > R(x_0, s_0) e^{\frac{-\Delta}{\rho}}$$

où $R(x_0, s_0) \stackrel{def}{=} e^{\frac{\nu \ln < s_0, x_0 > + F(x_0) + F_*(s_0) - \nu \ln \nu + \nu}{\rho}}$.

□

Corollaire 7.1 *Supposons que $\rho = \gamma\sqrt{\nu}$ où γ désigne une constante supérieure ou égale à 1 et d'autre part, que nous disposions d'un point strictement admissible du primal x_0, d'un point strictement admissible du dual (y_0, s_0) ainsi que d'un algorithme qui, à chaque itération, remplace l'itéré primal-dual strictement admissible (x_k, s_k) par le couple (x_{k+1}, s_{k+1}) de même nature et ce, de telle façon que*

$$\phi(x_{k+1}, s_{k+1}) \leq \phi(x_k, s_k) - \delta$$

où δ est strictement positif.
Alors, quel que soit le réel ϵ strictement positif considéré, nous atteignons l'ϵ-optimalité primale-duale après K itérations, où

$$K = \lceil \frac{\gamma\sqrt{\nu}}{\delta}[(1 + \frac{\sqrt{\nu}}{\gamma}) \ln < s_0, x_0 > + \frac{F(x_0) + F_*(s_0)}{\gamma\sqrt{\nu}} - \frac{\sqrt{\nu}}{\gamma} \ln \nu + \frac{\sqrt{\nu}}{\gamma} + \ln \frac{1}{\epsilon}] \rceil.$$

<u>Preuve:</u>
Nous devons montrer l'assertion suivante :

$$< s_K, x_K > \leq \epsilon.$$

Remarquons, tout d'abord, que l'hypothèse relative à la diminution de la fonction potentiel primale-duale implique l'inégalité ci-dessous :

$$\phi(x_K, s_K) \leq \phi(x_0, s_0) - K\delta.$$

Nous appliquons alors le Théorème 7.1 avec les identifications $\rho = \gamma\sqrt{\nu}$, $x = x_K$, $s = s_K$ et $\Delta = K\delta$. Nous obtenons de la sorte l'inégalité suivante :

$$< s_K, x_K > \leq < s_0, x_0 > e^{\frac{\nu \ln < s_0, x_0 > + F(x_0) + F_*(s_0) - \nu \ln \nu + \nu}{\gamma\sqrt{\nu}}} e^{\frac{-K\delta}{\gamma\sqrt{\nu}}}.$$

La définition de K nous permet de majorer le second membre de celle-ci. La borne ainsi obtenue se réduit, après simplification, à ϵ.

□

Le Théorème 7.1 montre qu'une diminution de la fonction potentiel primale-duale entraîne un rapprochement de l'optimalité primale-duale. Notre extension de l'algorithme de Kojima, Mizuno et Yoshise doit donc s'efforcer de diminuer la valeur de sa fonction potentiel au cours de ses itérations.

Le Corollaire 7.1 exploite ce résultat afin d'établir les conditions à satisfaire pour atteindre l'ϵ-optimalité primale-duale ainsi que le nombre d'itérations nécessaires à l'obtention de celle-ci. Il nous apprend, en effet, qu'il suffit que la fonction potentiel primale-duale décroisse à chaque itération d'une constante pour obtenir, après $O(\sqrt{\nu}\ln\frac{1}{\epsilon})$ itérations, une solution primale-duale ϵ-optimale. Comme précédemment, nous avons, lors de la détermination de la borne de complexité, négligé des quantités. A nouveau, deux raisons ont motivé cette négligence. D'une part, les quantités dont nous n'avons pas tenu compte, dépendent des problèmes (P) et (D) considérés et de l'initialisation. D'autre part, nous supposons ϵ pris suffisamment petit que pour rendre ces quantités négligeables par rapport à $\frac{\gamma\sqrt{\nu}}{\delta}\ln\frac{1}{\epsilon}$.

Par ailleurs, nous remarquons que la borne de complexité associée à notre méthode de réduction de potentiel primale-duale est plus fine que celle des méthodes de réduction de potentiel primales décrites dans le chapitre précédent.

7.2 Mise en oeuvre de procédures assurant la satisfaction des conditions de convergence

Suposons, comme dans le Corollaire 7.1, qu'à chaque itération de notre généralisation de la méthode de Kojima, Mizuno et Yoshise, nous disposions d'un point strictement admissible du primal $\hat{x}$ et d'un point strictement admissible du dual $(\hat{y}, \hat{s})$. La question à laquelle nous répondons dans ce point est celle de la détermination du prochain itéré primal-dual de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 7.1 et le Corollaire 7.1.

7.2.1 Détermination des directions de descente

La satisfaction des conditions de convergence nécessite une diminution constante de la fonction potentiel primale-duale à chaque itération. D'autre part, nous voulons obtenir un algorithme primal-dual symétrique. Dès lors, nous devons, à chaque itération, déterminer une direction de descente primale et une direction de descente duale de la fonction potentiel primale-duale. Nous détaillons, dans la suite de cette subdivision, l'obtention de celles-ci.

Nous calculons, tout d'abord, une direction de descente de $\phi(.,s)$ en $(\hat{x}, \hat{s})$ que nous désignons par $-\Delta\hat{x}$. Nous recherchons celle-ci dans le noyau de A. La présence de cette contrainte s'explique par notre volonté de générer un prochain itéré primal qui soit un point strictement admissible du problème primal (P). Il n'est pas nécessaire d'introduire la contrainte primale d'appartenance à l'intérieur de K puisque le respect de celle-ci est pris en charge par la fonction barrière F. Par ailleurs, nour préférons à la méthode de Newton une méthode de quasi-Newton. Celle-ci nous autorise à approximer la dérivée seconde de ϕ par rapport à x en $(\hat{x}, \hat{s})$ par le hessien de F évalué en un point quelconque de l'intérieur de K. Ce hessien présente l'avantage d'être défini positif. Dès lors, le problème de minimisation sous contrainte suivant :

$$\begin{array}{ll} \min\limits_{p} & < u, p > + \frac{1}{2} < F''(\omega)p, p > \\ sc. & Ap = 0 \end{array} \tag{7.1}$$

où $u \stackrel{def}{=} \phi'_x(\hat{x}, \hat{s})$ et où ω désigne le point "scaling" primal associé à $(\hat{x}, \hat{s})$, est convexe. Remarquons que l'existence et l'unicité de ω sont garanties par le Théorème 2.2. Le lecteur pourrait s'étonner du fait que nous ayons approximé la dérivée seconde de ϕ par rapport à x en $(\hat{x}, \hat{s})$ non pas par le hessien de F en $\hat{x}$ mais par le hessien de F en ω. Ce choix n'est évidemment pas anodin. Il se justifie, en fait, par notre volonté d'obtention d'une méthode primale-duale qui soit symétrique. Nous expliciterons davantage cette justification dans le point 7.4. La convexité tant de l'ensemble admissible que de la fonction objectif du problème (7.1) nous permet de le réécrire, de façon équivalente, sous la forme d'un système de Karush-Kühn-Tücker. De la sorte, la direction de recherche $\Delta\hat{x}$ est obtenue en résolvant le système linéaire en $(-\lambda, \Delta\hat{x})$ suivant :

$$\begin{array}{rcl} -A^*\lambda + F''(\omega)\Delta\hat{x} & = & u \\ A\Delta\hat{x} & = & 0 \end{array} \tag{7.2}$$

où λ désigne le vecteur des multiplicateurs de Lagrange associé à la contrainte d'égalité $Ap = 0$. L'existence et l'unicité de la direction de recherche $\Delta\hat{x}$ sont garanties par la Proposition 5.1 puisque, comme annoncé, nous constatons que la direction de recherche générée s'apparente à une projection oblique. De plus, cette constatation nous autorise à ne pas aborder la question de la résolution du système (7.2).

Nous déterminons, ensuite, une direction de descente de $\phi(x,.)$ en $(\hat{x}, \hat{s})$ que nous désignons par $-\Delta\hat{s}$. Nous voulons générer un nouvel itéré dual $(\hat{y} - \alpha\Delta\hat{y}, \hat{s} - \alpha\Delta\hat{s})$ qui soit un point strictement admissible du problème dual (D). Dès lors, nous recherchons des directions de recherche $\Delta\hat{y}$ et $\Delta\hat{s}$ telles que $A^*\Delta\hat{y} + \Delta\hat{s}$ soit nul. Par ailleurs, il n'est pas nécessaire d'introduire la contrainte duale d'appartenance à l'intérieur de K^* imposée aux variables d'écart. Le respect de celle-ci est, en effet, pris en charge par la fonction barrière conjuguée F_*. A nouveau, nour préférons à la méthode de Newton une méthode de quasi-Newton. Celle-ci nous autorise à approximer la dérivée seconde de ϕ par rapport à s en $(\hat{x}, \hat{s})$ par le hessien de F_* évalué en un point quelconque de l'intérieur de K^*. Ce hessien présente l'avantage d'être défini positif. Dès lors, le problème de minimisation sous contrainte suivant :

$$\begin{array}{ll} \min\limits_{r,q} & < v, q > + \frac{1}{2} < q, F_*''(t) q > \\ sc. & A^* r + q = 0 \end{array} \tag{7.3}$$

où $v \stackrel{def}{=} \phi_s'(\hat{x}, \hat{s})$ et où t désigne le point "scaling" dual associé à $(\hat{s}, \hat{x})$, est convexe. Remarquons que l'existence et l'unicité de t sont garanties par le Corollaire 2.1. Le lecteur pourrait, comme précédemment, s'étonner du fait que nous ayons approximé la dérivée seconde de ϕ par rapport à s en $(\hat{x}, \hat{s})$ non pas par le hessien de F_* en $\hat{s}$ mais par le hessien de F_* en t. De nouveau, ce choix est délibéré et s'explique, comme nous le soulignerons dans le point 7.4, par notre volonté d'obtention d'une méthode primale-duale qui soit symétrique. La convexité tant de l'ensemble admissible que de la fonction objectif du problème (7.3) nous permet de le réécrire, de façon équivalente, sous la forme d'un système de Karush-Kühn-Tücker. De la sorte, les directions de recherche $\Delta\hat{y}$ et $\Delta\hat{s}$ sont obtenues en résolvant le système linéaire en $(\mu, \Delta\hat{y}, \Delta\hat{s})$ suivant :

$$\begin{array}{rcl} A\mu & = & 0 \\ A^*\Delta\hat{y} + \Delta\hat{s} & = & 0 \\ \mu + F_*''(t)\Delta\hat{s} & = & v \end{array} \tag{7.4}$$

où μ désigne le vecteur des multiplicateurs de Lagrange associé à la contrainte d'égalité $A^*r + p = 0$.

Nous constatons que la connaissance d'une direction de descente primale et d'une direction de descente duale de la fonction potentiel primale-duale nécessite la résolution de deux systèmes linéaires; ce qui, au point de vue calculatoire, est relativement lourd. Heureusement, l'exploitation du caractère "ν-self-scaled" de la barrière F associée au cône K permet de réduire à un le nombre de systèmes à résoudre pour obtenir ces directions. La fin de cette subdivision est consacrée à l'explicitation de cette réduction. Le système (7.2) peut être réécrit, de façon équivalente, comme suit :

$$\begin{array}{rcl} A\Delta\hat{x} & = & 0 \\ A^*\lambda + \gamma & = & 0 \\ \gamma + F''(\omega)\Delta\hat{x} & = & u \end{array} \tag{7.5}$$

où $u \stackrel{def}{=} \phi_x'(\hat{x}, \hat{s}) = \frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{s} + F'(\hat{x})$. Puisque le système (7.2) s'apparente à une définition de projections obliques, la Proposition 5.1 garantit l'unicité et l'existence de λ et de $\Delta\hat{x}$. L'unicité de λ couplée à l'injectivité de l'opérateur A^* suffit à assurer l'unicité de γ. Dès lors, le système (7.5) en $(\Delta\hat{x}, \lambda, \gamma)$ admet une solution unique. Nous considérons alors le système linéaire (7.4) et nous nous attachons à réexprimer, de façon équivalente, sa troisième équation. Nous pré-multiplions les deux membres de celle-ci par l'inverse de $F_*''(t)$ et obtenons, de la sorte, l'égalité suivante :

$$[F_*''(t)]^{-1}\mu + \Delta\hat{s} = [F_*''(t)]^{-1}v.$$

L'utilisation de la relation liant t à ω, établie dans les commentaires relatifs au Corollaire 2.1, et de la deuxième égalité décrite (1.9) conduit à l'égalité ci-dessous :

$$[F_*''(t)]^{-1} = F''(\omega).$$

La troisième équation de (7.4) se réécrit donc comme suit :

$$F''(\omega)\mu + \Delta\hat{s} = F''(\omega)v,$$

où $v \stackrel{def}{=} \phi_s'(\hat{x}, \hat{s}) = \frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{x} + F_*'(\hat{s})$. Les définitions de v, de ω et de u ainsi que la relation (2.12) justifient le développement ci-dessous :

$$\begin{aligned} F''(\omega)v &= \frac{\nu+\rho}{<\hat{s},\hat{x}>}F''(\omega)\hat{x} + F''(\omega)F_*'(\hat{s}) \\ &= \frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{s} + F'(\hat{x}) \\ &= u. \end{aligned}$$

Par conséquent, nous obtenons la réécriture suivante de la troisième équation du système (7.4) :

$$F''(\omega)\mu + \Delta\hat{s} = u.$$

Le système (7.4) est donc équivalent au système en $(\mu, \Delta\hat{y}, \Delta\hat{s})$ suivant :

$$\begin{aligned} A\mu &= 0 \\ A^*\Delta\hat{y} + \Delta\hat{s} &= 0 \\ F''(\omega)\mu + \Delta\hat{s} &= u. \end{aligned}$$

Or, ce dernier coïncide avec le système (7.5) dont nous avons établi l'existence et l'unicité de la solution. Les égalités suivantes sont dès lors satisfaites :

$$\begin{aligned} \mu &= \Delta\hat{x}, \\ \Delta\hat{y} &= \lambda, \\ \Delta\hat{s} &= \gamma. \end{aligned}$$

Il est donc suffisant pour déterminer une direction de descente primale et une direction de descente duale de la fonction potentiel primale-duale de résoudre l'unique système linéaire suivant :

$$\begin{aligned} A\Delta\hat{x} &= 0 \\ A^*\Delta\hat{y} + \Delta\hat{s} &= 0 \\ F''(\omega)\Delta\hat{x} + \Delta\hat{s} &= u. \end{aligned} \tag{7.6}$$

La similitude entre les systèmes (7.5) et (7.6) étend au système (7.6) la validité des commentaires relatifs au système (7.5). Dès lors, l'existence et l'unicité des directions de recherche sont assurées. D'autre part, la méthode de calcul des projections obliques décrite à la fin du Chapitre 5 permet la détermination de $\Delta\hat{x}$ et de $\Delta\hat{s}$. $\Delta\hat{y}$ est, quant à lui, obtenu par la deuxième équation de (7.6). Soulignons, finalement, le rôle non négligeable tenu par le caractère "ν-self-scaled" de la barrière F associée au cône K dans la diminution du coût calculatoire inhérent à la détermination des directions de recherche. Cette diminution justifie déjà en partie le recours aux points "scaling" primal et dual dans les méthodes de quasi-Newton définissant les directions de descente primale et duale.

7.2.2 Diminution constante de la fonction potentiel

Théorème 7.2 *Etant donnés $\hat{x}$ l'itéré primal actuel et $(\hat{y}, \hat{s})$ l'itéré dual actuel, nous déterminons les directions de recherche $\Delta\hat{x}$, $\Delta\hat{y}$ et $\Delta\hat{s}$ comme décrit ci-dessus. Alors nous pouvons trouver un α strictement positif tel que*

(i.)

$$x_+ \stackrel{def}{=} \hat{x} - \alpha\Delta\hat{x} \quad \textit{est un point strictement admissible du primal}$$
$$(y_+, s_+) \stackrel{def}{=} (\hat{y} - \alpha\Delta\hat{y}, \hat{s} - \alpha\Delta\hat{s}) \quad \textit{est un point strictement admissible du dual}$$

(i.)

$$\phi(x_+, s_+) - \phi(\hat{x}, \hat{s}) \leq -\frac{\sqrt{3}}{2} + \ln(1 + \frac{\sqrt{3}}{2})$$

Preuve:

1. Nous nous intéressons à la stricte admissibilité primale-duale du triplet (x_+, y_+, s_+). Les directions de recherche $\Delta\hat{x}$, $\Delta\hat{y}$ et $\Delta\hat{s}$ ont été définies de telle façon que les contraintes primale et duale d'égalité soient satisfaites et ce, quelle que soit la valeur prise par α. La contrainte d'appartenance de x_+ à l'intérieur de K est vérifiée lorsque α appartient à l'intervalle $[0, \frac{1}{\sigma_{\hat{x}}(\Delta\hat{x})})$ et celle de s_+ à l'intérieur de K^* lorsque α appartient à l'intervalle $[0, \frac{1}{\sigma^*_{\hat{s}}(\Delta\hat{s})})$. Nous recherchons, dès lors, α parmi les réels strictement inférieurs à $\frac{1}{\bar{\sigma}}$ où

$$\bar{\sigma} \stackrel{def}{=} \max\{\sigma_{\hat{x}}(\Delta\hat{x}), \sigma^*_{\hat{s}}(\Delta\hat{s})\}.$$

2. Nous définissons la quantité μ par

$$\mu \stackrel{def}{=} \|u\|^*_\omega$$

et nous développons μ^2 de deux façons différentes. La première aboutit à une minoration de μ^2. Les définitions de u et de la norme $\|.\|_\omega^*$ ainsi que l'hypothèse de symétrie imposée aux hessiens de F justifient le développement suivant :

$$\begin{aligned}
\mu^2 &= (\|\frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{s}+F'(\hat{x})\|_\omega^*)^2 \\
&= <\frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{s}+F'(\hat{x}),[F''(\omega)]^{-1}(\frac{\nu+\rho}{<\hat{s},\hat{x}>}\hat{s}+F'(\hat{x}))> \\
&= \frac{(\nu+\rho)^2}{<\hat{s},\hat{x}>^2}<\hat{s},[F''(\omega)]^{-1}\hat{s}> +2\frac{\nu+\rho}{<\hat{s},\hat{x}>}<F'(\hat{x}),[F''(\omega)]^{-1}\hat{s}> \\
&\quad + <F'(\hat{x}),[F''(\omega)]^{-1}F'(\hat{x})> .
\end{aligned}$$

La définition de ω ainsi que le Théorème 2.2 impliquent les deux égalités ci-dessous :

$$\begin{aligned}
\hat{x} &= [F''(\omega)]^{-1}\hat{s} \\
F_*'(\hat{s}) &= [F''(\omega)]^{-1}F'(\hat{x})
\end{aligned}$$

L'emploi de celles-ci et de l'assertion (1.3) dans notre développement de μ^2 conduit à la suite d'égalités suivantes :

$$\begin{aligned}
\mu^2 &= \frac{(\nu+\rho)^2}{<\hat{s},\hat{x}>}+2\frac{\nu+\rho}{<\hat{s},\hat{x}>}<F'(\hat{x}),\hat{x}>+<F'(\hat{x}),F_*'(\hat{s})> \\
&= \frac{(\nu+\rho)^2-2\nu(\nu+\rho)}{<\hat{s},\hat{x}>}+<F'(\hat{x}),F_*'(\hat{s})> \\
&= <F'(\hat{x}),F_*'(\hat{s})> -\frac{\nu^2-\rho^2}{<\hat{s},\hat{x}>}. \qquad (7.7)
\end{aligned}$$

La contrainte imposée à ρ lors de la définition de la fonction potentiel primale-duale permet de majorer $\nu^2-\rho^2$ par $\nu(\nu-1)$ et donc, d'obtenir la minoration suivante de μ^2 :

$$\mu^2 \geq <F'(\hat{x}),F_*'(\hat{s})> -\frac{\nu(\nu-1)}{<\hat{s},\hat{x}>}. \qquad (7.8)$$

Nous développons, ensuite, μ^2 par une autre méthode. Celle-ci se base sur le système (7.6) qui génère les directions de recherche $\Delta\hat{x}$, $\Delta\hat{s}$ et $\Delta\hat{y}$. La troisième équation de celui-ci fournit, en effet, une expression de u. La première conclut, quant à elle, à l'appartenance de $\Delta\hat{x}$ au noyau de A et la seconde à l'appartenance de $\Delta\hat{s}$ à l'image de A^*. Ces deux dernières appartenances et l'orthogonalité des deux sous-espaces vectoriels qui y sont impliqués entraînent l'annulation du produit scalaire $<\Delta\hat{s},\Delta\hat{x}>$. Dès lors, nous obtenons, grâce également à l'hypothèse de symétrie imposée aux hessiens de F, le développement suivant de μ^2 :

$$\begin{aligned}
\mu^2 &= (\|F''(\omega)\Delta\hat{x}+\Delta\hat{s}\|_\omega^*)^2 \\
&= <F''(\omega)\Delta\hat{x}+\Delta\hat{s},[F''(\omega)]^{-1}(F''(\omega)\Delta\hat{x}+\Delta\hat{s})> \\
&= \|\Delta\hat{x}\|_\omega^2+2<\Delta\hat{s},\Delta\hat{x}> +(\|\Delta\hat{s}\|_\omega^*)^2 \\
&= \|\Delta\hat{x}\|_\omega^2+(\|\Delta\hat{s}\|_\omega^*)^2. \qquad (7.9)
\end{aligned}$$

3. Nous débutons la majoration de la différence $\phi(\hat{x}-\alpha\Delta\hat{x},\hat{s}-\alpha\Delta\hat{s})-\phi(\hat{x},\hat{s})$ que nous désignons par $\Delta\phi(\alpha)$ afin d'alléger les notations. Dans ce but, nous explicitons tout d'abord la quantité $<\hat{s}-\alpha\Delta\hat{s},\hat{x}-\alpha\Delta\hat{x}> - <\hat{s},\hat{x}>$. La nullité de $<\Delta\hat{s},\Delta\hat{x}>$ établie ci-dessus, la troisième équation de (7.6), la définition de ω ainsi que l'hypothèse de symétrie imposée aux hessiens de F expliquent les égalités suivantes :

$$\begin{aligned}
<\hat{s}-\alpha\Delta\hat{s},\hat{x}-\alpha\Delta\hat{x}> - <\hat{s},\hat{x}> &= -\alpha(<\Delta\hat{s},\hat{x}> + <\hat{s},\Delta\hat{x}>)\\
&= -\alpha(<u-F''(\omega)\Delta\hat{x},\hat{x}> + <\hat{s},\Delta\hat{x}>)\\
&= -\alpha(<u,\hat{x}> - <F''(\omega)\hat{x},\Delta\hat{x}> + <\hat{s},\Delta\hat{x}>)\\
&= -\alpha <u,\hat{x}>.
\end{aligned}$$

Par ailleurs, la définition de u et la propriété (1.3) impliquent l'égalité ci-dessous :

$$<u,\hat{x}>=\rho.$$

Dès lors, nous aboutissons à la relation suivante :

$$<\hat{s}-\alpha\Delta\hat{s},\hat{x}-\alpha\Delta\hat{x}> - <\hat{s},\hat{x}>=-\alpha\rho.$$

Nous majorons, ensuite, $\Delta\phi(\alpha)$. La définition de la fonction potentiel primale-duale et l'égalité démontrée ci-dessus justifient les égalités suivantes :

$$\begin{aligned}
\Delta\phi(\alpha) &= (\nu+\rho)\ln<\hat{s}-\alpha\Delta\hat{s},\hat{x}-\alpha\Delta\hat{x}> +F(\hat{x}-\alpha\Delta\hat{x})+F_*(\hat{s}-\alpha\Delta\hat{s})\\
&\quad -(\nu+\rho)\ln<\hat{s},\hat{x}> -F(\hat{x})-F_*(\hat{s})\\
&= (\nu+\rho)\ln\Big(\frac{<\hat{s},\hat{x}> + <\hat{s}-\alpha\Delta\hat{s},\hat{x}-\alpha\Delta\hat{x}> - <\hat{s},\hat{x}>}{<\hat{s},\hat{x}>}\Big)\\
&\quad +F(\hat{x}-\alpha\Delta\hat{x})-F(\hat{x})+F_*(\hat{s}-\alpha\Delta\hat{s})-F_*(\hat{s})\\
&= (\nu+\rho)\ln\Big(1-\frac{\alpha\rho}{<\hat{s},\hat{x}>}\Big)+F(\hat{x}-\alpha\Delta\hat{x})-F(\hat{x})+F_*(\hat{s}-\alpha\Delta\hat{s})-F_*(\hat{s}).
\end{aligned}$$

Le Théorème 3.2 nous permet alors de majorer le terme rendant compte de la variation de la barrière "ν-self-scaled" F, pour tout α appartenant à $[0,\frac{1}{\sigma_{\hat{x}}(\Delta\hat{x})})$. De plus, la Proposition 2.4 valide l'application de ce même théorème à la barrière conjuguée F_*. Dès lors, nous pouvons également majorer le terme rendant compte de la variation de la barrière conjuguée F_* pour tout α appartenant à $[0,\frac{1}{\sigma^*_{\hat{s}}(\Delta\hat{s})})$. Nous obtenons, de la sorte, la majoration suivante de $\Delta\phi(\alpha)$:

$$\begin{aligned}
\Delta\phi(\alpha) &\leq (\nu+\rho)\ln\Big(1-\frac{\alpha\rho}{<\hat{s},\hat{x}>}\Big)\\
&\quad -\alpha<F'(\hat{x}),\Delta\hat{x}> +\frac{\|\Delta\hat{x}\|^2_{\hat{x}}}{\sigma^2_{\hat{x}}(\Delta\hat{x})}(-\alpha\sigma_{\hat{x}}(\Delta\hat{x})-\ln(1-\alpha\sigma_{\hat{x}}(\Delta\hat{x})))\\
&\quad -\alpha<\Delta\hat{s},F'_*(\hat{s})> +\frac{(\|\Delta\hat{s}\|^*_{\hat{s}})^2}{(\sigma^*_{\hat{s}}(\Delta\hat{s}))^2}(-\alpha\sigma^*_{\hat{s}}(\Delta\hat{s})-\ln(1-\alpha\sigma^*_{\hat{s}}(\Delta\hat{s}))),
\end{aligned}$$

valable pour tout α appartenant à $[0, \frac{1}{\bar{\sigma}})$. Finalement, l'exploitation de la concavité du logarithme népérien nous conduit à l'inégalité ci-dessous :

$$\Delta\phi(\alpha) \leq -\alpha\Delta_0 + \Delta_{\hat{x}}(\alpha) + \Delta_{\hat{s}}(\alpha)$$

où

$$\begin{aligned}
\Delta_0 &\stackrel{def}{=} \frac{\rho(\nu+\rho)}{<\hat{s},\hat{x}>} + < F'(\hat{x}), \Delta\hat{x} > + < \Delta\hat{s}, F_*'(\hat{s}) >, \\
\Delta_{\hat{x}}(\alpha) &\stackrel{def}{=} \frac{\|\Delta\hat{x}\|_{\hat{x}}^2}{\sigma_{\hat{x}}^2(\Delta\hat{x})}(-\alpha\sigma_{\hat{x}}(\Delta\hat{x}) - \ln(1-\alpha\sigma_{\hat{x}}(\Delta\hat{x}))), \\
\Delta_{\hat{s}}(\alpha) &\stackrel{def}{=} \frac{(\|\Delta\hat{s}\|_{\hat{s}}^*)^2}{(\sigma_{\hat{s}}^*(\Delta\hat{s}))^2}(-\alpha\sigma_{\hat{s}}^*(\Delta\hat{s}) - \ln(1-\alpha\sigma_{\hat{s}}^*(\Delta\hat{s}))).
\end{aligned}$$

4. Nous interrompons momentanément notre majoration de $\Delta\phi(\alpha)$ afin d'estimer séparément Δ_0, $\Delta_{\hat{x}}(\alpha)$ et $\Delta_{\hat{s}}(\alpha)$. Nous débutons l'estimation de Δ_0 par l'explicitation de $< \Delta\hat{s}, F_*'(\hat{s}) >$. L'expression de $\Delta\hat{s}$ extraite de la troisième équation de (7.6) et la définition de u expliquent l'égalité ci-dessous :

$$< \Delta\hat{s}, F_*'(\hat{s}) >= \frac{\nu+\rho}{<\hat{s},\hat{x}>} < \hat{s}, F_*'(\hat{s}) > + < F'(\hat{x}), F_*'(\hat{s}) > - < F''(\omega)\Delta\hat{x}, F_*'(\hat{s}) > .$$

Les égalités décrites dans le Théorème 2.2, l'hypothèse de symétrie imposée aux hessiens de F et la propriété (1.3) permettent de réécrire le premier et le dernier produits scalaires intervenant dans le second membre de l'égalité ci-dessus comme suit :

$$\begin{aligned}
< \hat{s}, F_*'(\hat{s}) > &= < F''(\omega)\hat{x}, [F''(\omega)]^{-1}F'(\hat{x}) >= -\nu, \\
< F''(\omega)\Delta\hat{x}, F_*'(\hat{s}) > &= < F'(\hat{x}), \Delta\hat{x} > .
\end{aligned}$$

Cette égalité devient, dès lors, la suivante :

$$< \Delta\hat{s}, F_*'(\hat{s}) >= \frac{-\nu(\nu+\rho)}{<\hat{s},\hat{x}>} + < F'(\hat{x}), F_*'(\hat{s}) > - < F'(\hat{x}), \Delta\hat{x} > .$$

Par conséquent, nous pouvons réécrire Δ_0 de la façon suivante :

$$\begin{aligned}
\Delta_0 &= \frac{\rho(\nu+\rho)}{<\hat{s},\hat{x}>} - \frac{\nu(\nu+\rho)}{<\hat{s},\hat{x}>} + < F'(\hat{x}), F_*'(\hat{s}) > \\
&= < F'(\hat{x}), F_*'(\hat{s}) > - \frac{\nu^2-\rho^2}{<\hat{s},\hat{x}>} \\
&= \mu^2.
\end{aligned}$$

La relation (7.7) justifie la dernière égalité du développement ci-dessus. Nous nous attachons ensuite à majorer $\Delta_{\hat{x}}(\alpha)$ et $\Delta_{\hat{s}}(\alpha)$. Pour ce faire, nous établissons quelques inégalités utiles. Le point (*ii.*) du Corollaire 3.1 implique la relation ci-dessous :

$$F''(\hat{x}) \leq \sigma_{\hat{x}}^2(\omega)F''(\omega)$$

et donc, par définition des inégalités relatives au cône des opérateurs semi-définis positifs, l'inégalité suivante :

$$\|\Delta\hat{x}\|_{\hat{x}}^2 \leq \sigma_{\hat{x}}^2(\omega)\|\Delta\hat{x}\|_{\omega}^2. \tag{7.10}$$

Par conséquent, nous obtenons, grâce également à la propriété (3.11) et à l'expression de μ^2 décrite en (7.9), les inégalités ci-dessous :

$$\sigma_{\hat{x}}(\Delta\hat{x}) \leq \|\Delta\hat{x}\|_{\hat{x}} \leq \sigma_{\hat{x}}(\omega)\|\Delta\hat{x}\|_{\omega} \leq \sigma_{\hat{x}}(\omega)\mu. \tag{7.11}$$

Par ailleurs, remarquons l'égalité entre $\sigma_{\hat{x}}(\omega)$ et $\sigma_{\hat{s}}^*(t)$. Celle-ci s'explique par la nature de ω, par la relation liant t à ω, par la première égalité décrite en (1.2) ainsi que par le point (*iii.*) du Théorème 2.1. Nous désignons par σ cette valeur commune. La Proposition 2.4 valide l'application du point (*ii.*) du Corollaire 3.1 au cône dual et à sa barrière associée. De la sorte, nous obtenons la relation ci-dessous :

$$F_*''(\hat{s}) \leq \sigma^2 F_*''(t)$$

et donc, par définition des inégalités relatives au cône des opérateurs semi-définis positifs, l'inégalité suivante :

$$(\|\Delta\hat{s}\|_{\hat{s}}^*)^2 \leq \sigma^2(\|\Delta\hat{s}\|_t^*)^2.$$

Or $\|\Delta\hat{s}\|_t^* = \|\Delta\hat{s}\|_{\omega}^*$. En effet, la seconde égalité décrite en (1.9) et la relation liant t à ω entraînent l'égalité ci-dessous :

$$[F''(\omega)]^{-1} = F_*''(t).$$

Par conséquent, nous obtenons, grâce également aux relations (3.11) et (7.9), les inégalités suivantes :

$$\sigma_{\hat{s}}^*(\Delta\hat{s}) \leq \|\Delta\hat{s}\|_{\hat{s}}^* \leq \sigma\|\Delta\hat{s}\|_t^* = \sigma\|\Delta\hat{s}\|_{\omega}^* \leq \sigma\mu. \tag{7.12}$$

Nous déduisons, finalement, des suites d'inégalités (7.11) et (7.12) la majoration suivante de $\bar{\sigma}$:

$$\bar{\sigma} \leq \sigma\mu. \tag{7.13}$$

Nous majorons $\Delta_{\hat{x}}(\alpha)$ grâce à l'inégalité (7.10) et obtenons, de la sorte, l'inégalité ci-dessous :

$$\Delta_{\hat{x}}(\alpha) \leq \frac{\sigma^2\|\Delta\hat{x}\|_{\omega}^2}{\sigma_{\hat{x}}^2(\Delta\hat{x})}(-\alpha\sigma_{\hat{x}}(\Delta\hat{x}) - \ln(1 - \alpha\sigma_{\hat{x}}(\Delta\hat{x}))).$$

La définition de $\bar{\sigma}$ et l'appartenance supposée de α à $[0, \frac{1}{\bar{\sigma}})$ entraînent que :

$$\alpha\sigma_{\hat{x}}(\Delta\hat{x}) \leq \alpha\bar{\sigma} < 1.$$

Dès lors, le point (*iii.*) de la Proposition 3.2 garantit la satisfaction de l'inégalité ci-dessous :

$$\frac{-\alpha\sigma_{\hat{x}}(\Delta\hat{x}) - \ln(1-\alpha\sigma_{\hat{x}}(\Delta\hat{x}))}{\alpha^2\sigma_{\hat{x}}^2(\Delta\hat{x})} \leq \frac{-\alpha\bar{\sigma} - \ln(1-\alpha\bar{\sigma})}{\alpha^2\bar{\sigma}^2}.$$

Nous obtenons donc la majoration suivante de $\Delta_{\hat{x}}(\alpha)$:

$$\Delta_{\hat{x}}(\alpha) \leq \frac{\sigma^2\|\Delta\hat{x}\|_\omega^2}{\bar{\sigma}^2}(-\alpha\bar{\sigma} - \ln(1-\alpha\bar{\sigma})).$$

Un raisonnement similaire conduit à la majoration suivante de $\Delta_{\hat{s}}(\alpha)$:

$$\Delta_{\hat{s}}(\alpha) \leq \frac{\sigma^2(\|\Delta\hat{s}\|_\omega^*)^2}{\bar{\sigma}^2}(-\alpha\bar{\sigma} - \ln(1-\alpha\bar{\sigma})).$$

5. Nous achevons la majoration de $\Delta\phi(\alpha)$. La réécriture de Δ_0, les bornes obtenues sur $\Delta_{\hat{x}}(\alpha)$ et $\Delta_{\hat{s}}(\alpha)$ ainsi que l'égalité (7.9) conduisent à l'inégalité suivante :

$$\Delta\phi(\alpha) \leq -\alpha\mu^2 + \frac{\sigma^2\mu^2}{\bar{\sigma}^2}(-\alpha\bar{\sigma} - \ln(1-\alpha\bar{\sigma})),$$

dont le second membre peut être réécrit comme suit :

$$-\alpha\sigma^2\mu^2\left(\frac{1}{\sigma^2} + \frac{1}{\bar{\sigma}}\right) - \frac{\sigma^2\mu^2}{\bar{\sigma}^2}\ln(1-\alpha\bar{\sigma}).$$

On remarque que l'attribution de la valeur suivante :

$$\frac{1}{\sigma^2 + \bar{\sigma}}$$

à α minimise l'expression ci-dessus. Soulignons également la stricte positivité de σ^2. Celle-ci permet à $\alpha = \frac{1}{\sigma^2+\bar{\sigma}}$ de satisfaire la contrainte d'appartenance à $[0, \frac{1}{\bar{\sigma}})$ qui valide les applications antérieures du Théorème 3.2 et du point (*iii.*) de la Proposition 3.2 et qui garantit l'appartenance de x_+ et de s_+ respectivement aux intérieurs de K et de K^*. Le remplacement de α par $\frac{1}{\sigma^2+\bar{\sigma}}$ dans la majoration de $\Delta\phi(\alpha)$ obtenue, mène à l'inégalité suivante :

$$\Delta\phi(\alpha) \leq \frac{-\mu^2}{\bar{\sigma}} + \frac{\sigma^2\mu^2}{\bar{\sigma}^2}\ln\left(1 + \frac{\bar{\sigma}}{\sigma^2}\right).$$

On montre facilement que le second membre de celle-ci peut être réécrit, de façon équivalente, comme suit :

$$\frac{-\mu^2}{\sigma^2}\left[\frac{\tau - \ln(1+\tau)}{\tau^2}\right],$$

où $\tau \stackrel{def}{=} \frac{\bar{\sigma}}{\sigma^2}$. Remarquons que τ est strictement plus grand que -1. D'autre part, l'inégalité (7.13) donne lieu à la relation ci-dessous :

$$\tau \leq \frac{\mu}{\sigma}.$$

Par conséquent, le point (*ii.*) de la Proposition 3.2 garantit la satisfaction de l'inégalité suivante :

$$\frac{\tau - \ln(1+\tau)}{\tau^2} \geq \frac{\frac{\mu}{\sigma} - \ln(1+\frac{\mu}{\sigma})}{\frac{\mu^2}{\sigma^2}}.$$

Ainsi, nous obtenons la borne suivante sur $\Delta\phi(\alpha)$:

$$\Delta\phi(\alpha) \leq -(\frac{\mu}{\sigma} - \ln(1+\frac{\mu}{\sigma})).$$

Par ailleurs, l'inégalité (7.8) et le Théorème 4.2 impliquent que :

$$\frac{\mu^2}{\sigma^2} \geq \frac{3}{4}.$$

Cette inégalité couplée à la croissance du logarithme népérien permet d'aboutir au résultat escompté, à savoir :

$$\Delta\phi(\alpha) \leq -(\frac{\sqrt{3}}{2} - \ln(1+\frac{\sqrt{3}}{2})).$$

□

Rappelons que l'objectif de ce point est, étant donnés les itérés actuels $\hat{x}$ et $(\hat{y}, \hat{s})$ points strictement admissibles respectivement du primal et du dual, de déterminer le prochain itéré primal-dual de telle sorte que soient satisfaites les conditions de convergence établies par le Théorème 7.1 et le Corollaire 7.1. Dans la subdivision précédente, nous avons explicité l'obtention d'une direction de descente primale $-\Delta\hat{x}$ et d'une direction de descente duale $-\Delta\hat{s}$ de la fonction potentiel primale-duale ϕ en l'itéré actuel $(\hat{x}, \hat{s})$. Nous avons également dégagé la direction de recherche $-\Delta\hat{y}$ en les variables duales. La satisfaction des conditions de convergence nécessite encore la garantie de l'existence de longueurs de pas dans les directions $-\Delta\hat{x}$, $-\Delta\hat{y}$ et $-\Delta\hat{s}$ pour lesquelles une diminution constante de la fonction potentiel primale-duale ϕ et la stricte admissibilité primale-duale du prochain itéré sont assurées. Le Théorème 7.2 comble ce manquement. Il garantit, en effet, l'existence d'une longueur de pas α, commune aux trois directions de recherche, qui satisfait les exigences mentionnées ci-dessus.

7.3 Algorithme

1. **Initialisation:** choisir x_0 un point strictement admissible du primal et (y_0, s_0) un point strictement admissible du dual.

2. **k-ième itération** $(k \geq 0)$ **:**

   (a) détermination des directions de recherche:

   - détermination du point "scaling" primal associé à (x_k, s_k) et noté ω_k;
   - calcul de u_k:

   $$u_k = \frac{\nu + \rho}{< s_k, x_k >} s_k + F'(x_k);$$

   - résolution du système linéaire suivant:

   $$\begin{aligned} A\Delta x_k &= 0 \\ A^* \Delta y_k + \Delta s_k &= 0 \\ F''(\omega_k)\Delta x_k + \Delta s_k &= u_k. \end{aligned}$$

   (b) détermination de la longueur de pas:

   - calcul de $\sigma_{x_k}(\Delta x_k)$, de $\sigma^*_{s_k}(\Delta s_k)$ et de $\sigma_k = \sigma_{x_k}(\omega_k)$;
   - détermination de $\bar{\sigma}_k$:

   $$\bar{\sigma}_k = \max\{\sigma_{x_k}(\Delta x_k), \sigma^*_{s_k}(\Delta s_k)\};$$

   - calcul de la longueur de pas initiale:

   $$\bar{\alpha}_k = \frac{1}{\sigma_k^2 + \bar{\sigma}_k};$$

   - détermination de la longueur de pas α_k telle que
     - $\alpha_k \in [0, \bar{\alpha}_k]$
     - $\phi(x_k - \alpha_k \Delta x_k, s_k - \alpha_k \Delta s_k) \leq \phi(x_k - \bar{\alpha}_k \Delta x_k, s_k - \bar{\alpha}_k \Delta s_k)$.

   (c) détermination de l'itéré suivant:

   $$\begin{aligned} x_{k+1} &= x_k - \alpha_k \Delta x_k \\ s_{k+1} &= s_k - \alpha_k \Delta s_k \\ y_{k+1} &= y_k - \alpha_k \Delta y_k. \end{aligned}$$

Fin de l'itération

7.4 Appréciation de la méthode obtenue

Notre objectif était d'obtenir une méthode de réduction de potentiel qui résout simultanément et symétriquement les problèmes (P) et (D). L'extension obtenue atteint cet objectif. En effet, nous avons montré dans la preuve du Théorème 7.2 que σ_k peut indifféremment être défini par $\sigma_{x_k}(\omega_k)$ et par $\sigma^*_{s_k}(t_k)$. Ce fait établit la symétrie primale-duale de la longueur de pas utilisée. D'autre part, la troisième équation du système générant les directions de recherche est équivalente à la suivante:

$$\Delta x_k + F''_*(t_k)\Delta s_k = v_k, \tag{7.14}$$

où $v_k \stackrel{def}{=} \phi'_s(x_k, s_k)$. Cette équivalence est garantie par la réécriture de la troisième équation du système (7.4) décrite à la fin de la subdivision relative à la détermination des directions de descente. Celle-ci nous assure de la symétrie primale-duale des directions de descente primale et duale générées. Rappelons que la réécriture de la troisième équation du système (7.4) sur laquelle se base cette équivalence exploite les relations existant entre les points "scaling" primal et dual associés à un même couple de points intérieurs de K et de K^*. Notre volonté d'obtention d'une méthode symétrique justifie dès lors l'intervention des points "scaling" primal et dual dans les approximations des dérivées secondes de la fonction potentiel primale-duale réalisées dans les méthodes de quasi-Newton définissant les directions de descente primale et duale.

Par conséquent, l'emploi des nouvelles distances introduites au Chapitre 3 et le caractère "ν-self-scaled" de la barrière associée au cône primal, qui garantit l'existence et l'unicité des points "scaling", permettent l'établissement d'un algorithme primal-dual qui soit symétrique.

Par ailleurs, l'extension obtenue fournit un algorithme à grands pas. En effet, la longueur de pas utilisée ne confine pas le prochain itéré primal et le prochain itéré dual respectivement à l'intérieur des boules unité centrées en x_k et en s_k, et définies par les normes locales primale et duale. Cette longueur de pas excède effectivement $\frac{1}{\|\Delta x_k\|_{\hat{x}}}$ et $\frac{1}{\|\Delta s_k\|^*_{\hat{s}}}$. A nouveau, la grandeur de la longueur de pas s'explique par la restriction aux barrières et cônes "self-scaled" ainsi que par l'introduction des nouvelles distances réalisées au Chapitre 3. Ces deux facteurs valident en effet les bornes sur les variations de la barrière F et de sa conjuguée F_* pour une longueur de pas α supérieure à $\frac{1}{\|\Delta x_k\|_{\hat{x}}}$ et $\frac{1}{\|\Delta s_k\|^*_{\hat{s}}}$.

Ces commentaires concluent donc à la pertinence de la considération des barrières "ν-self-scaled" et des nouvelles distances introduites au Chapitre 3 dans les méthodes de réduction de potentiel primales-duales symétriques à grands pas.

Nous clôturons cette appréciation par deux constatations. Si ρ est de la forme $\gamma\sqrt{\nu}$

où γ désigne une constante supérieure ou égale à 1, la méthode de réduction de potentiel primale-duale détaillée ci-dessus atteint son ϵ-optimalité après $O(\sqrt{\nu}\ln\frac{1}{\epsilon})$ itérations tandis que les méthodes de réduction de potentiel primales décrites au chapitre précédent n'atteignaient leur ϵ-optimalité qu'après $O(\nu\ln\frac{1}{\epsilon})$ itérations. Nous constatons donc une convergence plus rapide de la méthode de réduction de potentiel primale-duale. D'autre part, les directions de recherche primale et duale s'apparentent, comme annoncé, à des projections obliques. En effet, Δx_k est la projection de $\phi'_x(x_k, s_k)$ sur le noyau de A relativement à $F''(\omega_k)$. Nous remplaçons la troisième équation du système définissant les directions de recherche par l'équation équivalente (7.14). La considération du système ainsi obtenu permet de percevoir Δs_k comme la projection de $\phi'_s(x_k, s_k)$ sur l'image de A^* respectivement à $F''_*(t_k)$.

Conclusion

Le but de ce mémoire était d'étudier la résolution du problème de programmation convexe conique "self-scaled" au moyen de méthodes de réduction de potentiel. L'étude en question devait notamment souligner l'impact de l'hypothèse suivant laquelle le cône associé au problème considéré est "self-scaled". Nous avons ainsi montré que, sous cette hypothèse, l'extension à la programmation convexe conique de méthodes de réduction de potentiel primales et primale-duale symétrique utilisées en programmation linéaire, donne lieu à des algorithmes à grands pas.

Il aurait également été intéressant d'apprécier les répercussions de l'introduction de l'hypothèse de "self-scaling" sur des méthodes de suivi de chemin applicables à la programmation convexe conique. Le lecteur désireux d'approfondir ce sujet, peut entre autres se référer à [1, §9] et [10] .

Bibliographie

[1] YU. NESTEROV and M.J. TODD. Self-scaled barriers and interior-point methods in convex programming. *Mathematics of Operations Research*, 22:1-42, 1997.

[2] YU. NESTEROV and A. NEMIROVSKII. Interior Point Polynomial Algorithms in Convex Programming. *SIAM*, Philadelphia, 1994.

[3] N.K. KARMARKAR. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373-395, 1984.

[4] C.C. GONZAGA. Polynomial affine algorithms for linear programming. *Mathematical Programming*, 49:7-21, 1990.

[5] M. KOJIMA, S. MIZUNO and A. YOSHISE. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementary problems. *Mathematical Programming*, 50:331-342, 1991.

[6] R.T. ROCKAFELLAR. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

[7] I. EKELAND and R. TEMAM. *Convex Analysis and Variational Problems.* North-Holland, Amsterdam, 1976.

[8] M.J. TODD. Potential-reduction methods in mathematical programming. *Mathematical Programming*, 76:3-45, 1996.

[9] YU. NESTEROV. *Long-step strategies in interior point potential-reduction methods.* Departement SES COMIN, Université de Genève, 1993.

[10] YU. NESTEROV and M.J. TODD. Primal-Dual Interior-Point Methods for Self-Scaled Cones. Core Discussion paper 9544, Center for Operations Research and Econometrics, Université Catholique de Louvain, 1995.

Errata

page de présentation : Dans le titre de l'abstract, lire *convex* au lieu de *convexe.*
page 19 : Dans le développement de $\ln(\rho^2 - \|s\|^2)$ en fin de page, lire $\ln[\frac{4\tau^2}{(\tau^2-\|x\|^2)^2} - \frac{4\|x\|^2}{(\tau^2-\|x\|^2)^2}]$ au lieu de $\ln[\frac{4\tau^2}{(\tau^2-\|x\|^2)^2} - \frac{4|x\|^2}{(\tau^2-\|x\|^2)^2}]$.
page 23 : Dans le point (*iii.*) en milieu de page, lire $F_*(\hat{s}) \geq F_*(F''(v)\,z) + < \hat{s} - F''(v)\,z, F'_*(F''(v)\,z) >$ au lieu de $F_*(\hat{s}) \geq F_*(F''(v)\,z) + < \hat{s} - F''(v)\,z, F'(F''(v)\,z) >$.
page 25 : Dans le dernier paragraphe, lire *caractère "self-scaled"* au lieu de *caractère "ν-self-scaled".*
page 27 : Dans la réécriture de la borne inférieure de $\phi(v)$, lire $< s, v > +F(v) - F(x) + \nu$ au lieu de $< s, v > +F(v) - F(x) - \nu$.
page 27 : Dans la définition de la fonction ψ, lire $< s, v > +F(v) - F(x) + \nu$ au lieu de $< s, v > +F(v) - F(x) - \nu$.
page 29 : Dans l'égalité (2.14), lire $F'_*(s)$ au lieu de $F'_*(x)$.
page 30 : Dans le point (*i.*) du Corollaire 2.2, lire *Quels que soient v un élément de K et x un point intérieur de K* au lieu de *Quels que soient v et x des points intérieurs de K.*
page 45 : En fin de preuve, lire *la relation (1.1)* au lieu de *la relation (2.2).*
page 52 : Dans le second membre de la troisième équation décrite en (4.4), lire $< F''(\omega + \alpha v)v, v >$ au lieu de $< F'(\omega + \alpha v)v, v >$.
page 64 : Dans la deuxième phrase du dernier paragraphe, lire *pourront s'appliquer à ces directions de recherche* au lieu de *pourront s'appliquer ces directions de recherche.*
page 71 : En fin de page dans le second membre de l'inégalité explicitant la relation $\phi(\bar{x};\xi) \leq \phi(x_0;\xi_0) - \Delta$, lire $\nu \ln < c - \xi_0 d, x_0 > +F(x_0) - \Delta$ au lieu de $\nu \ln < c - \xi d, x_0 > +F(x_0) - \Delta$.
page 74 : A la fin du deuxième paragraphe, lire $\frac{\nu}{\delta}\ln(\frac{1}{\epsilon})$ au lieu de $\frac{\nu}{\rho}\ln(\frac{1}{\epsilon})$.
page 82 : Dans l'avant dernière expression mathématique, lire p_+ au lieu de $p - +$.
page 87 : Dans la première phrase du dernier paragraphe, lire *Corollaire 6.2* au lieu de *Corolloraire 6.2.*
page 88 : Dans le second membre de la dernière égalité, lire $\frac{\mu}{<c,\hat{x}>-\xi}c + F'(\hat{x})$ au lieu de $\frac{\mu}{<c,\hat{x}>-\xi} + F'(\hat{x})$.
page 95 : Dans le point (*ii.*) du Théorème 6.4, lire $\max^2\{p_+, \|p\|_{\hat{x}}\}$ au lieu de $\max^2\{p_+, \|p\|^2_{\hat{x}}\}$.
page 97 : Dans les seconds membres des deux premières inégalités, lire $-\alpha < F'(\hat{x}), p >$ au lieu de $-\alpha < F'(x), p >$.
page 103 : Dans la dernière phrase, lire *grâce à l'hypothèse* au lieu de *grâce l'hypothèse.*
page 109 : Dans le second point (*i.*) du Théorème 7.2, lire (*ii.*) au lieu de (*i.*).